REPORT RESUMES

ED 015 864 24 RE 001 079
FACTORS AFFECTING LEARNING TO READ. FINAL REPORT.
BY- HAYES, ROBERT B. WUEST, RICHARD C.
PENNSYLVANIA STATE DEPT. OF PUBLIC INSTRUCTION
REPORT NUMBER BR-6-1752 PUB DATE OCT 67
NEW CASTLE AREA SCHOOLS, PA.
GRANT OEG-1-7-06172-0299
EDRS PRICE MF-$0.75 HC-$6.68 165P.

DESCRIPTORS- *READING RESEARCH, READING ACHIEVEMENT, *LONGITUDINAL STUDIES, *GRADE 2, *GRADE 3, SILENT READING, ORAL READING, *TEACHING METHODS, PHONICS, BASIC READING, INITIAL TEACHING ALPHABET,

FOR THE THIRD YEAR, THE FOLLOWING FOUR DIFFERENT METHODS OF TEACHING READING WERE CONTRASTED--(1) THE INITIAL TEACHING ALPHABET, (2) A PHONIC, FILMSTRIP, WHOLE-CLASS APPROACH, (3) A WHOLE-WORD, ECLECTIC BASAL READER METHOD, AND (4) THE PRECEDING APPROACH SUPPLEMENTED BY A PHONICS PROGRAM. SOME 400 FIRST-GRADE PUPILS WERE RANDOMLY ASSIGNED TO TREATMENTS UPON ENTRANCE INTO FIRST GRADE AND WERE FOLLOWED AS INTACT CLASSES INTO SECOND AND THIRD GRADES WITH THE SAME METHODS, EXCEPT THAT I/T/A PUPILS TRANSITIONED INTO THE MERRILL "TREASURY OF LITERATURE" PROGRAM IN SECOND GRADE. A SMALL REPLICATIVE STUDY WAS DONE IN THE LAST 2 YEARS OF THE PROJECT. CONSIDERABLE INSERVICE EDUCATION WAS PROVIDED ALL TEACHERS. TEACHING WAS OBSERVED FREQUENTLY BY A FULL-TIME FIELD DIRECTOR AND OTHER SUPERVISORY PERSONNEL TO CHECK ON TEACHER COMPETENCE AND ADHERENCE TO METHOD. TEACHERS COMPLETED ACTIVITY LOGS AS AN ADDITIONAL METHOD SAFEGUARD. SIGNIFICANT DIFFERENCES RESULTED FROM ANALYSIS OF VARIANCE AND COVARIANCE WHICH GENERALLY WERE IN FAVOR OF PROGRAMS (1) AND (2). THIS SUGGESTED THAT A COMBINATION OF (1) AND (2) WOULD BE WORTHY OF FUTURE USE AND ADDITIONAL STUDY. PROGRAM (2) PRODUCED THE BEST OVERALL RESULTS, PARTICULARLY FOR THE HIGH IQ THIRD. HOWEVER, THE PUPIL RETENTION RATE WITH THIS PROGRAM SUGGESTED THAT IT MIGHT BE A BETTER PROGRAM WITH ABILITY GROUPING. (AUTHORS)

ED015864

FINAL REPORT
Cooperative Research Project No. 6-1752 - 24
USOE Contract No. OEG-1-7-06172-0299

FACTORS AFFECTING LEARNING TO READ

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

October 1967

RE001 079

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FACTORS AFFECTING LEARNING TO READ

Cooperative Research Project No. 6-1752
USOE Contract No. OEG-1-7-06172-0299

Robert B. Hayes
Director, Research Administration
and Coordination
Commonwealth of Pennsylvania
Harrisburg, Pennsylvania

Richard C. Wuest
Field Director
New Castle Area Schools
New Castle, Pennsylvania

October 1967

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

INTRODUCTION 1

METHOD 5

RESULTS 19

DISCUSSION 27

CONCLUSIONS 28

IMPLICATIONS 30

SUMMARY 31

REFERENCES 34

APPENDIXES

A. Hayes Teacher Rating Scale A-1

B. Grade III Teacher Log B-1

C. Grade II Teacher Log C-1

D. Statistical Data and Analysis - Grade II D-1

E. Statistical Data and Analysis - Grade III E-1

F. Pupil Retentions - A Study by Linda Nash F-1

ACKNOWLEDGMENTS

Grateful appreciation is extended to the following participating teachers of the New Castle Area Schools: Mrs. Karen Atkinson, Miss Mary Ann Barto, Mrs. Carol Baughman, Mrs. Joan Biasucci, Mrs. Margaretta Bratshie, Miss Kathleen Chiafullo, Mrs. Frances C. Christy, Mrs. Murl L. Cowher, Mrs. Alverta E. Crissman, Mrs. Diane DeEulio, Miss Sara D. Fields, Mrs. Florence Flannery, Miss Ellen E. Gallagher, Miss Margaretta E. Hess, Miss Margaret E. Kauffman, Mrs. Carol Ledwith, Mrs. Helen E. Lucas, Miss Lucile W. Lutton, Mrs. Judy Matthews, Mrs. Sue Michael, Mrs. Eleanor Miller, Mrs. C. Elise Murdoch, Mrs. Frances Nicholas, Mrs. Dorothy Purdue, Mrs. Vera S. Reed, Mrs. Freda Sampson, Mrs. Caroll Smith, Mrs. Helen B. Smith, Mrs. Nancy A. Smith, Mrs. Agnes W. Williams, Mrs. Ruth M. Wilson, and Mrs. Edna Mae Woodring.

Sincere appreciation is also extended to: Mr. Russell L. Horchler, Superintendent of Schools; Mr. Calvin DiCarlo, Assistant Superintendent of Schools; Mr. Eugene DeCaprio, School Psychologist; Mrs. Mary Elizabeth Morrison and Mr. William J. Nicholas, Guidance Counselors; Mr. William Thompson, Business Manager; and the following elementary school principals: Mr. John N. Cornelius, Mr. Robert M. Cowher, Mr. John W. Ellefson, Mr. Pat J. George, Mr. Orlando Lucidore, and Mr. Arthur H. Walker.

The generous cooperation of the following consultants is warmly acknowledged: Miss Ednamae Bruggeman, Scott, Foresman Company; Miss Margaret A. Lennox, J. B. Lippincott Company; Mr. Glenn McCracken, J. B. Lippincott Company; Miss G. Margaret Wilson, Charles E. Merrill Books, Inc.; and Mrs. Elaine Wonsavage, American Education Publications.

Thanks are also due Mr. Joseph F. Haenn, Mr. Thomas J. Rookey and Mr. Russell E. Dusewicz, Department of Public Instruction, for their assistance with the statistical analysis. Special recognition is due Miss Linda E. Nash, Department of Public Instruction, who studied pupils who were retained at the ends of Grades I and II.

INTRODUCTION

THE PROBLEM

In September 1964 the United States Office of Education, through its Cooperative Research Program, initiated a coordinated series of studies across the country in an attempt to discover answers to many questions related to beginning reading instruction which have plagued parents and educators for many years. This series of studies, which has come to be known as the "First Grade Reading Studies", originally included about 30,000 children in twenty-seven individual studies, conducted in a variety of locations through the cooperation of many colleges, universities, state departments of education, and local school districts. Each of the studies has been unique in that none was a duplication of another, but all attempted to maintain the same controls and evaluative techniques. Not all of the original studies were extended beyond one year, but a number of them have continued for two or three years, and a few plan to go further.

New Castle, Pennsylvania, has been the site of one of this series of reading studies. A primary goal of the New Castle study has been to determine which of four different approaches to beginning reading instruction was the most effective, but several related questions were also investigated. This study was a three-year longitudinal study with a modified replication. During the first two years of this investigation, many significant differences were found among the treatment groups (21 and 22), but it was recognized that similar results over a longer period of time and during a replicative study would add validity to the findings. Therefore, the third year (1966-1967) of the New Castle study was designed to follow the first- and second-grade classes from the preceding year into grades II and III to determine whether previous findings would be supported.

OBJECTIVES

This project attempted to determine reading achievement and attitudes resulting from continued teaching with four different approaches: Scott, Foresman; Lippincott; American Education Publications; and i/t/a-Charles E. Merrill.

More specifically, the objectives were:

1. Which of these methods was best for children of different ability levels?

2. What were the teaching characteristics of teachers whose students achieved above their reading expectancy levels?

3. What were the teaching characteristics of teachers whose students maintained high interest in reading?

4. What were the effects of each method on spelling achievement?

RELATED RESEARCH

Most research in the field of reading instruction has been conducted in relatively short time spans of one year or less, and very few studies have been carefully replicated in an attempt to lend greater validity to reported findings.

Holt observes, "In almost every instance of a well-planned, carefully executed experiment involving teachers and students who engage in new educational enterprises, the initial results are positive.... At the conclusion of the initial experiments, the obvious inference is that all schools and all teachers should adopt the new and discard the old. But careful examination of the broad application of a new program all too often reveals that the bright promises have not been fulfilled. In the hands of less dedicated, less interested (or less well-paid) teachers the new program becomes undistinguished and even on occasion less effective than the old" (15, p.188).

Roma Gans has recently stated, "The data coming off the press daily which cite the pros and cons of experiments with new materials and approaches for teaching reading to beginners after only one or two years of experimenting are not adequate" (12, p.15). Harris and others have reported, "Too often comparative investigations in reading have had a duration of one year or less, and too often, rather broad generalizations have been inferred from the results of these short lived endeavors" (14, p.311). Nevertheless, in these days of increased concern for education, there are growing pressures from the public, from publishing companies, and even from school administrators and teachers to find and adopt the "best" way to teach reading. Many school systems face and succumb to the urge to "jump on various bandwagons" which offer the solution to educational problems which have burdened us for years.

According to Harris and Serwer, "Accumulated evidence is abundant on the need to study long-term as well as short-term results of teaching procedures" (20, p.98). The reasons for this need are many. Initial results favoring one approach may not be supported in subsequent years of study; some teaching techniques may affect long-range behavior of students in ways which might not be apparent at the conclusion of a short-term study; the effects of various beginning reading approaches on future success on other school subjects needs to be evaluated; and Hawthorne effects must be controlled before drawing valid conclusions from research findings.

One of the most serious and prevalent weaknesses in educational research is the lack of awareness or control of Hawthorne effects which have been described as "...a phenomenon characterized by an awareness on the part of the subjects of special treatment created by artificial conditions" (10, p.118), resulting in improved scores which would not occur without the awareness of special treatment. Through the use of a longitudinal approach in which the same teachers and students are included in the study for a period of years, the experimental conditions tend to become routine, and the halo surrounding the approaches being evaluated tends to disappear.

Another method of assessing unusual motivation due to possible Hawthorne effects is to replicate the research, with the same teachers in the same geographic areas, under conditions where special attention, help, and publicity is markedly reduced from the original study.

It appears that comparatively few reading method replicative studies have been made and that these have occurred outside their original settings. For example, the Greenman study (19), where significant differences in first-grade achievement favored children taught by a phonic approach, was conducted in Champaign, Illinois; the Sparks and Fay study (20), which reported significant differences in comprehension favoring the phonic group at the end of grades I and II but not in higher grades, was performed in Louisville, Kentucky; and the Kelley study (15), which compared a phonic approach with a basal reader approach, was done in Murphysboro, Kentucky. Gray commented on similar situations when he wrote, "...the results secured in given experiments have not been repeated under similar conditions to validate original findings" (13, p.1087).

"An increase in generalizability can also be sought in the replication of experiments. Beyond this, a special advantage is often gained in replicating, in an experiment on a variable not previously studied, one or more variables whose effects have already been assessed as significant in a previous experiment.... This scientifically healthy practice is almost routine in the physical sciences, but, unhappily, is the exception in educational experimentation" (18, pp.659-660). Kinsella expresses the same viewpoint in the statement, "Can you imagine the medical profession accepting a treatment for a certain disease because it produced promising results in one experiment...?", and he continues, "repetition of encouraging experiments is a rare event in many areas of educational research...We must do more of it." (17, pp.88-89).

In summary, related research points to the need for longitudinal and replicative reading studies to provide increased reliability and validity of results.

METHOD

RESTATEMENT OF THE PROBLEM

The primary purpose of this study was to follow the first- and second-grade students of 1965-1966 into Grades II and III, as intact classes, to determine the effects of continued instruction in four different approaches to reading instruction.

EXPERIMENTAL DESIGN

The reading achievement (and a series of related questions) of nineteen Grade III classes and twelve Grade II classes was studied through the application of Design 6 (Post-test-Only Control Group Design) as described by Campbell and Stanley (8, pp.195-197). This is the same design used during two previous years of study. There was a slight modification of Design 6 in that standardized achievement testing was done in the middle as well as at the end of the year.

The independent treatment variables for both Grades II and III were: (1) A basal reader program using materials published by Scott, Foresman and Company, 1962 edition; (2) a phonic program which utilized correlated filmstrips and published by the J. B. Lippincott Company, 1963 edition; (3) an eclectic, combination program which used the materials of the Scott, Foresman Company (No. 1 above) supplemented with phonic booklets (Phonics and Word Power) published by American Education Publications, Inc.; and (4) a language arts approach using the initial teaching alphabet as a medium, represented by the materials of i/t/a Publications, Inc., 1963 edition. Treatment variable number four (i/t/a) is the only one which has been changed. This was necessary because that program was conceived and designed as one to be used only for the initial teaching of reading. By the time most children completed the first grade, they had transferred to traditional orthography. For this reason, Dr. Albert J. Mazurkiewicz, co-author of the Early-to-Read i/t/a Program and consultant to the first-grade i/t/a classes during the 1964-1965 school year, recommended the Treasury of Literature Series, published by Charles E. Merrill Books, Inc., as appropriate materials to use following transition from i/t/a.

The dependent variables which were the same for both grades were: (1) The Stanford Achievement Test, Primary II, administered in January and May; (2) the San Diego County Inventory of Reading Attitude, given in April; (3) the number of books children read independently were sampled by recording this item for the month of February; (4) the Gates Word Pronunciation Test, administered in April to a subsample of the second- and third-grade populations; and (5) the Gilmore Oral Reading Test, given to the same subsamples immediately following the administration of the Gates. In June of Grade III the reading subtests of the Stanford Intermediate I also were administered.

The participating book companies again provided consultant services to aid the teachers in using their materials according to the methods advocated by the companies. In an effort to control one cause of Hawthorne effects, these services were to be provided only to the third-grade teachers. Second-grade teachers were to have been experienced in the methods they were teaching, through previous years of experience in the study when they would have received the advantage of classroom observations, and workshops provided by the consultants. However, the three i/t/a-Merrill teachers of second grade who had not met with the consultants from the Charles E. Merrill Company for the complete series of workshops during the 1965-1966 phase of the study, and one new second-grade Scott, Foresman teacher were permitted to participate in the August meeting and the first two meetings during the year.

In August 1966 all third-grade teachers participated in an appropriate six-hour workshop conducted by their book company consultant, who explained teaching philosophy and provided concrete suggestions and directions to help the teachers begin their instructional programs correctly. Several weeks after school opened, during the third week of September, the consultants returned to New Castle to observe their teachers teach reading for forty-five to fifty minutes. Following the classroom observations, after-school workshop meetings were held for sixty to ninety minutes. This procedure afforded the consultants an opportunity to make specific suggestions and to offer constructive criticism. The teachers were able to raise questions, to discuss common problems, and to share ideas. The same plan was followed during the months of November, January and March, and teachers were compensated at the rate of four dollars ($4.00) per hour for the time they spent in the workshop meetings.

The following people were the consultants to the various groups: Miss Ednamae Bruggeman for the Scott, Foresman Company; Miss Margaret A. Lennox, Educational Consultant, and Dr. S. Glenn McCracken for the J. B. Lippincott Company; Mrs. Elaine Wonsavage, American Education Publications, Inc.; and Miss G. Margaret Wilson, Charles E. Merrill Books, Inc. Since the teachers in treatment variable number three (Phonics and Word Power) actually used the Scott, Foresman program supplemented with the phonics booklets, Miss Bruggeman and Mrs. Wonsavage visited those classrooms together, and jointly conducted the workshop meetings.

All teachers included in the study were also visited for forty-five to fifty minutes, at random, twelve times by supervisory personnel to check on adherence to material and method limitations, and to evaluate teacher effectiveness. The time of these visitations was announced to the teachers on the morning of each visit. The field director visited and rated each teacher on the Hayes Teacher Rating Scale (Appendix A) seven times during the school year; the assistant to the superintendent followed the same procedure twice; and the building principals observed and evaluated their teachers three times.

In addition to the classroom visitations and teacher ratings, all teachers submitted logs to the field director as another method safeguard. Third-grade teachers were required to keep logs during alternate weeks (Appendix B). On these forms they summarized the objectives of each lesson, the skills taught, the materials used, the grouping procedures followed, and the time spent teaching reading for each day of the weeks when logs were required. Since almost all second-grade teachers had participated in the study during the previous year when logs were also submitted at the end of alternate weeks, they were only required during the current year to record a summary (Appendix C) of the materials used and grouping procedures followed at the end of each month. This variation in requirements was followed as a means of reducing Hawthorne effects in the replicative study.

Attempts to control Hawthorne effects were also made through rigidly controlling the attention received by all treatment groups; by not permitting visitors to the classrooms except for supervisory personnel; by using second-grade teachers who had previous experience in the study with the same materials and methods; by largely eliminating second-grade workshops and observations by book company consultants; and by restricting the number of local presentations related to the study.

The children were encouraged to read widely from books on their own interest and reading levels, and opportunities were provided for them to select such materials. For purposes of reading instruction, however, the teachers were restricted to using only those materials recommended or suggested by their consultants, and materials available through the participating publishing companies were provided.

The policy of the New Castle Area Schools required second-grade teachers to spend five hundred thirty minutes of each week teaching reading, while third-grade classes received an average of four hundred fifteen minutes per week of reading instruction. Time during which teacher and students were in direct contact for the purpose of teaching reading and using those materials which were recommended by the book company consultants defined the reading instructional period. Supportive activities such as content area reading, independent reading, spelling, and creative writing were not counted into the weekly time limitations.

All classes commenced instruction on September 12, 1966 and the final testing was begun on May 15, 1967.

STATISTICAL ANALYSIS

A 4x3 factorial analysis of variance and covariance (where appropriate) was performed. In this analysis, factor A consisted of four methods of teaching reading while factor B represented three levels of intelligence (high, average and low). The preceding analysis involved random casting out of cases to produce an equal number of cases per cell. This resulted in 15 cases per IQ level, 45 per treatment, and a total N of 180 in Grade II and also in Grade III. The Stanford paragraph meaning scores were also analyzed for all students by an unweighted means analysis (9, pp.241-244).

For the analysis of variance involving 180 cases per grade, a Tukey (a) multiple range test was employed to determine which differences between means were contributing to significant F ratios. When analysis of covariance produced significant F ratios, Winer's multiple F test (9, pp.592-599) was used to compare differences between each appropriate pair of means. The analysis of variance, covariance and correlation matrices were performed at the Computation Center of The Pennsylvania State University, University Park, Pennsylvania. The multiple range tests were calculated by research assistants and associates in the Bureau of Research Administration and Coordination of the Pennsylvania Department of Public Instruction, Harrisburg, Pennsylvania.

POPULATION

According to the policy of the local school district, all prospective first-grade students are given the Lee-Clark Reading Readiness Test during the month of May preceding school entry. All students were ranked according to the scores attained on this test and were then randomly assigned, using a table of random numbers (6), to the required number of classrooms and treatment groups by attendance areas. This procedure was followed in selecting the population for the original study, and was repeated during the next year for the replication.

The original study included five classrooms per treatment group until one first-grade Scott, Foresman teacher became ill and was lost, with her class, from the study. Therefore, in the third grade there were nineteen classrooms, five per treatment group except for the Scott, Foresman group which had only four. The replicative study has included three classrooms per treatment group during both years.

Only those students for whom complete data was available were included in the analyses. Many reasons account for student attrition including: moving, retentions, or absences during testing periods. At the end of the 1964-1965 school year, 365 Grade I children were in the original study. By the end of second grade, 302 students remained; and 264 children remained in the study at the end of third grade (SF - 50; Lipp - 69; PWP - 70; and i/t/a-Merr - 75). The replicative study, during the year 1965-1966, included 248 first-grade students; and 213 remained by the end of second grade (SF - 62; Lipp - 56; PWP - 47; i/t/a-Merr - 48).

In October, during the first grade of each year, the Pintner-Cunningham Primary Test of Intelligence, 1964 revision, was administered and scored by the school psychologist. The mean intelligence quotients for the third-grade treatment groups were: Scott, Foresman - 98.49; Lippincott - 98.58; Phonics and Word Power - 96.98; and i/t/a-Merrill - 97.96. The mean IQ's, by levels, of the various treatment groups were:

	SF	Lipp	PWP	i/t/a-Merr
High IQ	112.40	114.07	108.87	112.66
Average IQ	99.67	98.93	98.40	97.07
Low IQ	83.40	82.73	83.67	84.13

In the replicative study, the mean intelligence quotients for each of the second-grade treatment groups were: Scott, Foresman - 105.49; Lippincott - 101.76; Phonics and Word Power - 102.67; and i/t/a-Merrill - 101.27. The mean IQ's, by levels, of the second-grade treatment groups were:

	SF	Lipp	PWP	i/t/a-Merr
High IQ	119.20	112.80	117.20	114.40
Average IQ	105.20	101.93	101.27	100.13
Low IQ	92.07	90.53	89.53	89.27

The average numbers of pupils per third-grade classroom, according to treatment groups, were: Scott, Foresman - 19; Lippincott - 19; Phonics and Word Power - 23; and i/t/a - Merrill - 21. In second grade of the replicative study, the average numbers of students per classroom were: Scott, Foresman - 26; Lippincott - 20; Phonics and Word Power - 21; i/t/a-Merrill - 18. It should be noted that not all students in each classroom were included in the study. As project children moved or were retained, it became necessary for the principals to add nonstudy students. The data collected on those additional students has not been analyzed, but they were treated in every other way as though they were part of the population of the study.

TEACHER CHARACTERISTICS

Of the nineteen teachers included in the third-grade phase of this study, eight of them (four Scott, Foresman; three Lippincott; and one Phonics and Word Power) had participated previously. Each of the second-grade teachers, according to the design of the study, was to have had previous experience in the study. However, administrative problems prevented the local superintendent from complying with this requirement, and one Scott, Foresman teacher and one i/t/a-Merrill teacher failed to meet this qualification. In addition, a Scott, Foresman teacher became ill at midyear and was replaced by a teacher on her first permanent assignment.

The average age of all third-grade teachers was about 41½ years, with the following averages for each of the treatment groups: Scott, Foresman - 38.25 years; Lippincott - 47.8 years; Phonics and Word Power - 38.8 years; and i/t/a-Merrill - 40.4 years. The second-grade teachers averaged nearly 37½ years of age, and the treatment groups averaged: Scott, Foresman - 33 years; Lippincott - 46 years; Phonics and Word Power - 45.7 years; and i/t/a-Merrill - 25 years.

Third-grade teachers averaged a little more than fourteen years of previous teaching experience, ranging from none (one Phonics and Word Power teacher) to twenty-eight years (one Scott, Foresman and one Lippincott teacher). A comparison of the average previous teaching experience of the third-grade treatment groups follows: Scott, Foresman - 11.75 years; Lippincott - 24.6 years; Phonics and Word Power - 8.8 years; and i/t/a-Merrill - 10.6 years. Two second-grade teachers from the Scott, Foresman group, including the substitute, had no previous teaching experience (except for student teaching and substituting), and one teacher (the Scott, Foresman teacher who became ill at midyear) had forty years of previous experience. The average years of previous teaching experience for all second-grade teachers was nearly eleven years, and the treatment groups compared as follows: Scott, Foresman - 11.25 years; Lippincott - 11.66 years; Phonics and Word Power - 18 years; and i/t/a-Merrill - 2.6 years.

Sixteen of the nineteen third-grade teachers had taught Grade III previously, ranging from one year to twenty-one years, with an average of 7.4 years of third-grade teaching experience before the 1966-1967 school term. The Scott, Foresman group averaged 6.75 years; Lippincott teachers averaged 12.2 years; Phonics and Word Power teachers had taught third grade for an average of 3.2 years before being included in this study; and the i/t/a-Merrill group of teachers averaged 7.2 years of such experience. Only two of the second-grade teachers (both Scott, Foresman) had never taught Grade II previously. The others ranged from one year to twenty-four years of previous second-grade teaching experience, and averaged 6.71 years. The averages of the teachers of the four second-grade treatment groups follows: Scott, Foresman - 6.25 years; Lippincott - 8 years; Phonics and Word Power - 10.3 years; and i/t/a-Merrill - 2.3 years.

Ten of the third-grade teachers were married, one was widowed, one was divorced, and seven were single. The married third-grade teacher had from none (three teachers) to four children (one teacher). Only two second-grade teachers were single and all others were married. The number of children the married second-grade teachers had ranged from none (six teachers) to three (three teachers).

All third-grade teachers had earned at least a bachelors degree and four of them (one Scott, Foresman; two Lippincott; and one Phonics and Word Power) had been awarded masters degrees. Twelve of the fifteen Grade III teachers (two Scott, Foresman; three Lippincott; three Phonics and Word Power; and four i/t/a-Merrill) with bachelors degrees had taken an average of eleven additional credits, and all teachers, except three who had not taught long enough, were permanently certified. There were two second-grade teachers (one Lippincott and one Phonics and Word Power) who had not earned at least a bachelors degree, and one (Scott, Foresman) who had received a masters degree plus three credits. Of the second-grade teachers with a bachelors degree, eight had taken an average of slightly over eleven additional credits. Three second-grade teachers, who had not taught a sufficient number of years, were provisionally certified. All others had received permanent certification.

METHODS AND MATERIALS

The four treatment variables, used in both second and third grades of this study, are summarized below:

1. The first group, which used materials published by Scott, Foresman and Company (5), utilized a basal reader approach with what has been called a "whole-word" method, a generalization which is somewhat misleading. At the beginning of first grade, students learn a basic sight vocabulary through a variety of meaningful activities, but thereafter, a well-organized program of phonetic and structural word analysis skills, as well as other methods of word identification, is taught. In addition, the program emphasizes understanding as the reason and end result of all reading activities. Many opportunities are provided to develop comprehension abilities. Furthermore, the program includes activities designed to develop an appreciation of good literature. The authors recognized and expect children to develop at differing rates, and ability grouping is utilized as one means of meeting

individual differences. This treatment variable will be referred to in this report as "SF".

2. The second group used reading materials published by the J. B. Lippincott Company (4). This approach has been known as a "phonic" program, but once again, this identification is somewhat erroneous. In the first grade, instruction is begun by teaching in isolation the "auditory and visual recognition" of the five vowels, followed by "auditory and visual recognition" of five consonants. From that point, the stories are presented and composed almost entirely of words containing only letters and phonetic elements which have been previously taught. Some phonetically irregular words are introduced as sight words. As students gain proficiency in phonic analysis, structural analysis skills are also taught.

One of the essential differences between this program and the Scott, Foresman approach to beginning reading instruction, is that the co-authors of this series regard the act of reading as a decoding process. They reason that meaning resides in language and once a child is able to break the code he will be able to read with understanding. Nevertheless, certain activities in the workbooks and teacher manuals are included to evaluate and improve comprehension.

Another difference is that this approach utilized whole-class techniques of instruction, but some ability grouping was done on a limited basis, usually during afternoon reading instructional periods. With very few exceptions, the children in a particular grade received reading instruction from books designed for that grade level.

A unique feature of this approach is the use of correlated textfilms which accompany the basic texts. These filmstrips contain condensed versions of the lessons presented in the books and were used for motivation, evaluation, review, or reteaching. This treatment variable will be referred to in this report as "Lipp".

3. The third treatment group used a combination approach. The program of Scott, Foresman and Company (as described above) was used as the basic reading instructional program. This approach was supplemented with phonics workbooks published by American Education Publications, Inc. (3). These booklets were intended to strengthen and enrich basic programs by providing teachers with appropriate materials to use when it is necessary to reteach specific word analysis techniques, or to reinforce daily lessons presented through the basic program. The teachers followed ability grouping procedures and were directed to select levels and pages for use according to recognized needs of the students and the organization of the

basic texts. In addition to a wide variety of sequentially developed phonic analysis skills, each level of these booklets also contains exercises designed to develop structural analysis techniques. This treatment variable will be referred to in this report as "PWP".

4. The fourth treatment group used the initial teaching alphabet (7) as the medium of reading instruction in first grade. Dr. Albert Mazurkiewicz was then the consultant to this group and he recommended a supplementary literature series published by Charles E. Merrill Books, Inc. (2) for use following transition to traditional orthography. This series had not been intended by its publishers as a means of developing basic reading skills, but was designed to supplement a strong developmental program by providing reading materials of high literary value. It was precisely for this reason that Dr. Mazurkiewicz selected it. He wanted materials which would extend the children's interests in reading, and felt that the more typical basal readers would be inappropriate.

The teachers who comprised this group felt a need for a more highly structural program and, with the consultant provided by Charles E. Merrill Books, Inc., agreed to supplement the literature series with other materials available through the company. These were intended to provide the teachers with a means of teaching various word analysis and comprehension skills. The literature series continued to be used regularly and the other Merrill materials came to be regarded as the program for developing basic reading skills. The literature series was used in place of basal readers as a means of providing worthwhile stories and poems through which healthy attitudes and interests in reading could be fostered.

Ability grouping procedures were followed in the skills development phase of this reading program, but a wide variety of grouping techniques was used with the literature series. The groups which were established for these materials were dependent upon the objectives of each lesson. Sometimes the entire class read, discussed, or dramatized a story; sometimes specific interest groups were established. Art activities, panel discussions, oral reading, dramatizations, varied book reporting techniques, and other procedures were followed to aid in the development of literary appreciation. This treatment variable will be referred to in this report as "i/t/a-Merr".

The specific materials which were used in this study are listed below by treatment group and grade level. For those treatments which used grouping techniques, it should be recognized that the teachers of one grade used materials of other grade levels as required by the specific needs of the students.

Scott, Foresman (SF)

Grade II

- *Friends Old and New* (2^1)
- *More Friends Old and New* (2^2)
- *Think and Do Book* (2^1)
- *Think and Do Book* (2^2)
- *Wide Horizons* (2)
- *What Next?* (Part One)
- *What Next?* (Part Two)
- *My Little Pictionary*
- *My Second Pictionary*
- *My Practice Pad* (2^1)
- *My Practice Pad* (2^2)
- *Invitations to Personal Reading* (2)

Grade III

- *Roads to Follow* (3^1)
- *More Roads to Follow* (3^2)
- *Think and Do Book* (3^1)
- *Think and Do Book* (3^2)
- *Wide Horizons* (3)
- *New Tall Tales* (Part One)
- *New Tall Tales* (Part Two)
- *My Second Pictionary*
- *Thorndike-Barnhart Beginning Dictionary*
- *My Practice Pad* (3^1)
- *My Practice Pad* (3^2)
- *Invitations to Personal Reading* (3)

Lippincott (Lipp)

Grade II

- *Basic Reading* (2^1)
- *Basic Reading* (2^2)
- *Workbook* (2^1)
- *Workbook* (2^2)
- *Basic Reading Textfilm* (2^1)
- *Basic Reading Textfilm* (2^2)

Grade III

- *Basic Reading* (3^1)
- *Basic Reading* (3^2)
- *Workbook* (3^1)
- *Workbook* (3^2)
- *Basic Reading Textfilm* (3^1)
- *Basic Reading Textfilm* (3^2)
- *Basic Reading Phonics Guide*

Scott, Foresman
Plus Phonics and Word Power (PWP)

The materials published by Scott, Foresman and Company, and listed above, were provided to all classes in this treatment variable. In addition, the following materials published by American Education Publications, Inc., were used:

Grade II

- *Phonics and Word Power* Program 2, Book A
- *Phonics and Word Power* Program 2, Book B
- *Phonics and Word Power* Program 2, Book C

Grade III

- *Phonics and Word Power* Program 3, Book A
- *Phonics and Word Power* Program 3, Book B
- *Phonics and Word Power* Program 3, Book C

i/t/a-Merrill (i/t/a-Merr)

Grade II

Happiness Hill (2)
Nicky (2)
Red Deer the Indian Boy (2)
Universal Workbook in Reading (2)
Universal Workbook in Phonics (2A)
Universal Workbook in Phonics (2B)

Grade III

Treat Shop (3)
Uncle Funny Bunny (3)
Scottie and His Friends (3)
Universal Workbook in Reading (3)
Universal Workbook in Phonics (3A)
Universal Workbook in Phonics (3B)

EVALUATIVE TECHNIQUES

Stanford Achievement Test, Primary II, Forms Y and X

All second-grade reading study classes were given four subtests (Word Meaning, Paragraph Meaning, Spelling, and Word Study Skills) of this test (Form Y) on January 10 and 11, 1967. The same subtests were administered to third-grade classes on January 12 and 13, 1967. The entire battery (including Science and Social Studies, Language, Arithmetic Computation, and Arithmetic Concepts) was administered to all reading study classes during the week of May 15, 1967. On the preceding date Form X was used.

Classroom teachers were responsible for administering these tests, but a neutral professional person was in each room during the test administration to provide needed assistance and to check on adherence to time limitations and standardized directions. Several days prior to the dates for giving the tests, the field director met with all teachers and assistants to review recommended testing procedures, to stress the need for strictly adhering to time limits and the directions provided in accompanying manuals. The January tests were scored by competent third parties under the supervision of the field director, and in May the tests were machine scored.

Stanford Achievement Test, Intermediate I, Form W

A preliminary analysis of the results obtained by third-grade students in January on the Primary II level of the Stanford indicated that many of the children scored at or near the top of the test. This was particularly true of children in the high IQ third. It was the feeling of the principal investigator and the field director that even more children would attain scores at the top of the test in May, so plans were made to administer the Word Meaning, Paragraph Meaning, Spelling and Word Study Skills subtests of the Stanford, Intermediate I, to all third-grade classes in the reading study on June 1, 1967. Teachers again administered the tests with the help of professional assistants; they were scored by competent third parties under the supervision of the field director.

Gates Word Pronunciation Test

This test is a list of words which become progressively more difficult. Students are shown the words and are asked to pronounce them.

The field director of the study individually administered the test to a randomly selected sample of eighteen Grade III students per treatment, stratified by intelligence thirds, between the dates of April 17, 1967 and April 26, 1967.

Following a series of training meetings, the two local guidance counselors, who had been trained in the individual administration of various psychological tests, gave the test to a random sample of fifteen Grade II students per treatment group, stratified by IQ thirds. The second-grade sample was tested between April 28, 1967 and May 4, 1967, and each counselor administered the test to thirty randomly selected students.

Gilmore Oral Reading Test, Form A

During the same time periods given above for the Gates Word List, the Gilmore Oral was individually administered to the same second- and third-grade subsamples by the guidance counselors and the field director respectively. The counselors were also trained by the field director to properly administer and score this test. The Gilmore provides evaluations of oral reading accuracy, comprehension, and rate of reading.

ADDITIONAL EVALUATION

To determine the teaching characteristics of teachers whose students achieved above their reading expectancy levels, the procedures described below were followed.

Teaching characteristics of each teacher were determined by analyzing the results attained on the Hayes Teacher Rating Scale (Appendix A). Those students who achieved above their expected levels were identified by computing Bond and Tinker Expectancy Grade Scores (1, pp.76-80) according to the following formula: years in school x IQ + 1.0, and comparing the results with grade equivalent scores attained on the Word Meaning, Paragraph Meaning, and Word Study Skills sections of the Stanford

Achievement Test, Form X, which was administered in May.

A child was judged to be above his reading expectancy level if he achieved grade equivalent scores which were one-half grade level or more above his expected scores on at least two out of three reading subtests of the Stanford. Those classes which contained fifty percent or more students who qualified as achieving above expected levels were identified, and the teaching characteristics of those teachers were analyzed and compared with the characteristics of teachers whose classes did not qualify.

The teaching characteristics of teachers whose students maintained a high interest in reading were determined by the following procedures.

Student interest in reading was determined by administering the San Diego County Inventory of Reading Attitude (Appendix D) to all reading study students on April 12, 1967. This inventory is composed of twenty-five items, related to reading interests, which are read to the students by their teachers. The children indicate their feelings toward each item, after hearing it, by circling "Yes" or "No". A raw score of nineteen (stanine score of six) or better was considered indicative of a better than average interest in reading.

After determining which students had an above-average interest in reading according to San Diego results, those classrooms containing fifty percent or more reading study students who qualified were identified, and the teaching characteristics (Hayes Teacher Rating Scale) of those teachers were analyzed and compared with teachers of classes which did not qualify.

The effects of each of the reading instructional methods included in this study upon achievement in spelling were determined by analyzing and comparing the results attained on the spelling sections of the Stanford Tests.

Pupil retentions were examined by the case study approach.

SUMMARY OF PROCEDURES

This study was designed to follow the 1965-1966 first- and second-grade students, as intact classes, into second and third grades respectively, in an attempt to determine the effects of continued teaching in four different approaches to beginning reading instruction.

Nineteen third-grade classes (four Scott, Foresman; five Lippincott; five Scott, Foresman plus Phonics and Word Power; and five i/t/a-Merrill) which had participated in the study during two previous years were included. The replicative portion of the study included twelve second-grade classes (three per treatment method) which had also been included during the previous year.

The dependent variables for each grade were standardized silent reading achievement tests, number of books read, and a reading attitude inventory. In addition, subsamples of eighteen third-grade students and fifteen second-grade students per treatment group, randomly selected by IQ thirds, were individually administered tests of word recognition and oral reading achievement. Pupil retentions were examined by the case study approach.

Regular teacher logs and frequent classroom visitations by supervisory personnel were used to insure adherence to time, material, and method limitations. Third-grade teachers attended five workshop meetings conducted by book company consultants who also visited these classrooms four times during the year. Limited help was provided to second-grade teachers as one means of controlling for Hawthorne effects.

RESULTS

JANUARY OVERALL ACHIEVEMENT - GRADE II

These results were adjusted statistically for differences in intelligence and teacher effectiveness ratings. The word meaning scores (Table 1, Appendix D) were significantly higher for Lippincott compared to each of the other three programs, while i/t/a-Merrill was also significantly higher than Phonics and Word Power. The paragraph meaning scores were significantly higher for Lippincott (Table 3, Appendix D) compared to both Scott, Foresman and Phonics and Word Power, while i/t/a-Merrill was also significantly higher than Phonics and Word Power. For Word Study Skills (Table 5, Appendix D) both Lippincott and i/t/a-Merrill were significantly higher than the other two programs. The results for Spelling (Table 7, Appendix D) produced significantly higher scores for Lippincott compared to Scott, Foresman and Phonics and Word Power, while i/t/a-Merrill also was significantly higher than Phonics and Word Power.

JANUARY ACHIEVEMENT BY ABILITY LEVELS - GRADE II

For the high IQ third the significant results were: (1) For Word Meaning (Table 9, Appendix D) Lipp was higher than SF and PWP, while i/t/a-Merr was higher than PWP; (2) for Paragraph Meaning (Table 10, Appendix D) Lipp was higher than SF and PWP; (3) for Word Study Skills (Table 11, Appendix D) Lipp was higher than SF and PWP, while i/t/a-Merr was higher than PWP; and (4) for Spelling (Table 12, Appendix D) Lipp and i/t/a-Merr were higher than the other two programs.

For the average IQ third the significant results were: (1) For Word Meaning (Table 13, Appendix D), Lipp and i/t/a-Merr were higher than the other two programs; (2) for Paragraph Meaning (Table 14, Appendix D), i/t/a-Merr and Lipp were higher than PWP; (3) for Word Study Skills (Table 15, Appendix D) i/t/a-Merr and Lipp were higher than SF and PWP, and (4) for Spelling (Table 16, Appendix D) i/t/a-Merr and Lipp were higher than the other two programs.

For the low IQ third the significant results were: (1) For Paragraph Meaning (Table 18, Appendix D) Lipp was higher than i/t/a-Merr; and (2) for Word Study Skills (Table 19, Appendix D) Lipp was higher than SF.

As an additional safeguard, paragraph meaning scores were also compared by an unweighted means analysis for all 211 pupils tested in Grade II in January 1967, and the results (Tables 25, 26, 27, and 28, Appendix D) were practically identical to those obtained with a random sample of 180 of the 211 pupils.

END-OF-YEAR OVERALL ACHIEVEMENT - GRADE II

The Stanford Achievement Test results were adjusted statistically for difference in intelligence. The word meaning scores (Table 29, Appendix D) were significantly higher for Lipp compared to each of the other three programs. The paragraph meaning results (Table 31, Appendix D) were significantly higher for Lipp compared to SF and PWP, while i/t/a-Merr was significantly higher than PWP. The Word Study Skills scores (Table 33, Appendix D) were significantly higher for Lipp compared to each of the other three programs, while i/t/a-Merr was significantly higher than SF and PWP. The Spelling results (Table 35, Appendix D) were significantly higher for Lipp than each of the other three programs, while i/t/a-Merr was significantly higher than SF and PWP. The Language scores (Table 37, Appendix D) were significantly higher for Lipp than SF and PWP, while i/t/a-Merr was significantly higher than PWP.

No significant differences resulted for the subsample of 60 pupils on the Gates Word List or the Gilmore Oral. The grand mean for the Gates was 26.97 with a standard deviation of 8.0. For the Gilmore Accuracy, Comprehension and Rate, the grand means were 29.28, 20.83, and 96.45 with standard deviations of 9.84, 4.11 and 25.79 respectively.

END-OF-YEAR ACHIEVEMENT BY ABILITY LEVELS - GRADE II

For the high IQ third the significant differences were: (1) For Word Meaning (Table 39, Appendix D) Lipp over SF; (2) for Spelling (Table 42, Appendix D) Lipp over SF and PWP; and (3) for Language (Table 43, Appendix D) Lipp over PWP.

For the average IQ third the significant differences were: (1) For Paragraph Meaning (Table 45, Appendix D) i/t/a-Merr over PWP; (2) for Word Study Skills (Table 46, Appendix D) both Lipp and i/t/a-Merr over SF and PWP; and (3) for Spelling (Table 47, Appendix D) both Lipp and i/t/a-Merr over PWP.

For the low IQ third the significant differences were: (1) For Word Meaning (Table 49, Appendix D) Lipp and SF over i/t/a-Merr; and (2) for Paragraph Meaning (Table 50, Appendix D) Lipp over i/t/a-Merr.

READING INTERESTS - GRADE II

There were no significant differences among treatments in attitude of pupils toward reading as measured by the San Diego Attitude Scale. The grand mean for the preceding scale for all 180 pupils was 19.23 with a standard deviation of 3.80.

For number of books read in the month of February 1967 there were significant differences as follows: (1) PWP over both i/t/a-Merr and SF, and also Lipp over SF (Table 54, Appendix D); (2) for the high third, PWP over i/t/a-Merr and SF (Table 56, Appendix D); and (3) for the low IQ third, Lipp over i/t/a-Merr (Table 58, Appendix D).

The three highest coefficients of correlations obtained between number of books read and other study variables were: .34 with May Stanford Arithmetic Computation, .31 with April Gilmore Oral Rate, and .30 with May Stanford Word Meaning. The three highest coefficients of correlations obtained between the San Diego Attitude Scale and other study variables were: .28 with Teacher Effectiveness Ratings, .17 with number of books read, and .16 with May Stanford Spelling (See Table 63, Appendix D).

TEACHING CHARACTERISTICS AND READING ACHIEVEMENT - GRADE II

Three of the twelve second-grade classes contained a majority of reading study students who achieved at least one-half grade level above their reading expectancy scores (Table 59, Appendix D). A further analysis of the results reveals that slightly more than thirty-two percent of the total second-grade population achieved at least one-half grade level above predicted scores.

Table 60, Appendix D indicates that there was little difference, according to ratings received on the Hayes Teacher Rating Scale, between those teachers whose classes qualified as achieving above predicted levels and those whose classes did not qualify. The mean scores on each major category of the rating scale were "Above Average". However, great differences existed among the individual teachers.

TEACHING CHARACTERISTICS AND PUPIL ATTITUDE - GRADE II

Nine of the second-grade classes contained a majority of students who had above average attitudes toward reading according to results attained on the San Diego County Inventory of Reading Attitude (Table 61, Appendix D). A further analysis of the results reveals that about sixty-two percent of the total second-grade population were judged to have better than average attitudes toward reading.

According to Table 62, Appendix D, those teachers whose classes were judged to have better than average interest in reading received higher ratings on the Hayes Teacher Rating Scale than teachers whose classes did not qualify. Under each major category of the rating scale, teachers of children showing high interest in reading received "Above Average" ratings while other teachers were judged to be "About Average". Once again, large differences existed among teachers when individual ratings were considered.

CONTENT ACHIEVEMENT AND RELATIONSHIPS - GRADE II

Significant differences resulted between Lipp and i/t/a-Merr in May 1967 on the Stanford Arithmetic Computation Test. The difference in the Lippincott mean of 29.29 (3.2) and the i/t/a-Merr mean of 23.67 (2.8) was significant at the .01 level of confidence. Other differences on content scores were not significant (Arithmetic Concepts means ranged from 21.40 for i/t/a-Merr and 24.67 for Lipp, while Science and Social Studies means ranged from 18.02 for SF to 20.09 for Lipp).

Significant correlations (Table 63, Appendix D) ranging from .40 to .65 were obtained between general reading skills (Word Meaning, Paragraph Meaning and Word Study Skills) and content areas (Science and Social Studies, Arithmetic Computation, and Arithmetic Concepts).

JANUARY OVERALL ACHIEVEMENT - GRADE III

These results were adjusted statistically for differences in intelligence and teacher effectiveness ratings. The Word Meaning Scores (Table 64, Appendix E) were significantly higher for Lipp compared to PWP. There were no significant differences for Paragraph Meaning (Table 66, Appendix E). For Word Study Skills (Table 68, Appendix E), Lippincott was significantly higher than SF and PWP, while i/t/a-Merr was significantly higher than SF. The results for Spelling (Table 70, Appendix E) produced significantly higher scores for i/t/a-Merr compared to SF and PWP, while Lipp also was significantly higher than SF.

JANUARY ACHIEVEMENT BY ABILITY LEVELS - GRADE III

For the high IQ third there were no significant differences among treatments (Tables 72-75, Appendix E). For the average IQ third the significant differences were: (1) For Word Study

Skills (Table 78, Appendix E) Lipp over SF; and (2) for Spelling (Table 79, Appendix E) i/t/a-Merr and Lipp over SF. For the low IQ third there was a significant difference in Word Study Skills (Table 82, Appendix E) favoring Lipp over SF.

As an additional safeguard Paragraph Meaning scores were also compared by an unweighted means analysis for all 264 pupils tested in Grade III in January 1967. The results (Tables 84, 85, 86 and 87, Appendix E) were practically identical to those obtained with a random sample of 180 of the 264 pupils.

END-OF-YEAR OVERALL ACHIEVEMENT - GRADE III

Since in January 1967 at least one-fifth of the pupils in Grade III had scored close to the top of the Primary II Battery, Form Y of the Stanford Achievement Test, it was decided to test pupils not only on the Stanford Primary II, Form X during the week of May 15, 1967, but to also test pupils on the reading and spelling subtests of the Stanford Intermediate I Battery on June 1, 1967. Treatment IQ means were close together (98.58 for Lipp, 98.49 for SF, 97.96 for i/t/a-Merr, and 96.98 for PWP) and so were average teacher effectiveness ratings (15.67 for SF, 15.40 for Lipp, 15.18 for PWP and 14.40 for i/t/a-Merr). Statistical analysis of IQ and teacher effectiveness ratings proved to be far from significant (Tables 92 and 93, Appendix E).

Analysis of the May 1967 Stanford Achievement results indicated significant results only in Word Study Skills (Table 98, Appendix E); these results favored both i/t/a-Merr and Lipp compared to SF. Greater differentiation of pupil achievement was possible on the Stanford Intermediate I Battery and more significant differences resulted in June than in May 1967.

On June 1, 1967 of Grade III on the Stanford Intermediate I Battery, Lipp was significantly higher than both SF and PWP for Word Meaning, Paragraph Meaning, and Spelling (Tables 104, 106 and 108, Appendix E). At the same time, SF was significantly lower than each of the other three programs on Word Study Skills (Table 110, Appendix E).

For the subsample of 72 pupils randomly selected for individual testing, the significant differences were i/t/a-Merr over Lipp and SF for the Gilmore Oral Comprehension (Table 114, Appendix E).

END-OF-YEAR ACHIEVEMENT BY ABILITY LEVELS - GRADE III

For the high IQ third the significant differences were: (1) For Paragraph Meaning (Table 121, Appendix E) Lipp over PWP; (2) for Gilmore Accuracy (Table 124, Appendix E) both Lipp and i/t/a-Merr over SF; and (3) for Gilmore Comprehension (Table 125, Appendix E) both Lipp and i/t/a-Merr over PWP and SF.

For the average IQ third the one significant difference was in the area of Word Study Skills (Table 130, Appendix E) in favor of Lipp over SF. For the low IQ third the only significant difference was in Gilmore Comprehension (Table 141, Appendix E) in favor of i/t/a-Merr over Lipp.

READING INTERESTS - GRADE III

There were no significant differences among treatments in attitude of pupils toward reading (Tables 152-155, Appendix E) as measured by the San Diego Attitude Scale. Significant correlations were obtained between the San Diego Attitude Scale results and many of the other variables (See Table 174, Appendix E) with the highest of these correlations being .56 with the Gates Word List scores.

For the number of books read in the month of February 1967 there was a significant difference favoring both Lipp and SF over PWP and i/t/a-Merr (Table 157, Appendix E). For the high IQ third a significant difference in books read also resulted in favor of both Lipp and SF compared to i/t/a-Merr (Table 158, Appendix E). The three highest coefficients of correlations (Table 174, Appendix E) obtained between number of books read and other variables were: .44 with Gilmore Oral Rate, .33 with May Arithmetic Computation, and .30 with May Language.

TEACHING CHARACTERISTICS AND READING ACHIEVEMENT - GRADE III

Eight of the nineteen third-grade classes qualified as achieving above their reading expectancy levels according to results achieved on the Stanford Achievement Test, Primary II, Form X (Table 162, Appendix E). Of those eight classes, six also qualified at the end of Grades I and II. Also consistent with last year's results, nearly forty-seven percent of the total third-grade population was judged to be reading at least one-half grade level above predicted scores.

Table 163, Appendix E, indicates that teachers whose classes were judged to be achieving above their reading expectancy levels received "Superior" ratings under the category of Personality and "Above Average" ratings in each of the other categories. These results were slightly better than those achieved by the other teachers and were highly consistent with last year's results. Great variations among teachers existed when individual ratings were considered.

There were eleven third-grade classes in which fifty percent or more of the students achieved at least one-half grade level above their reading expectancy levels according to the results of the Stanford Achievement Test, Intermediate I, Form W (Table 164, Appendix E). Six of those eleven classes also qualified as over-achieving on the Primary II level of the Stanford in Grades II and III and on the Primary I at the end of the first grade. About fifty-nine percent of the Grade I pupils were judged to be over-achieving compared to forty-seven percent when the Primary II was used at the completion of second and third grade.

According to the results of the Hayes Teacher Rating Scale, the mean scores attained by teachers of classes in which a majority of the students achieved at least one-half grade level above expected scores were about the same as other teachers. All received "Above Average" mean ratings, and wide differences existed when individual ratings were compared.

TEACHING CHARACTERISTICS AND PUPIL ATTITUDE - GRADE III

Table 166, Appendix E, indicates that fourteen of the nineteen third-grade classes had better than average attitudes toward reading on the basis of their performance on the San Diego County Inventory of Reading Attitude. Ten of those classes were judged to have above average attitudes in second grade when there were thirteen such classes, and nine also qualified in Grade I. Over fifty-nine percent of the total third-grade population were judged to have better than average attitudes toward reading compared to about sixty-five percent who qualified last year.

The results indicated in Table 167, Appendix E, show that the teachers of students who were judged to have above average attitudes toward reading were rated about the same as teachers whose classes did not qualify. Their mean ratings were all in the "Above Average" range, but there were greater differences when individual ratings were compared.

CONTENT ACHIEVEMENT AND RELATIONSHIPS - GRADE III

No significant differences resulted among treatments in Science and Social Studies, Arithmetic Computation, and Arithmetic

Concepts as measured by the Primary II Battery of the Stanford Achievement Test (Tables 168, 170, and 172, Appendix E). Significant correlations (Table 174, Appendix E) ranging from .42 to .72 were obtained between general reading skills (Word Meaning, Paragraph Meaning and Word Study Skills) and content areas (Science and Social Studies, Arithmetic Computation, and Arithmetic Concepts). The results of the pupil retention study are reported in Appendix F. Here it is noted that retention is the answer when there is insufficient provision for individual diagnosis and appropriate personalized instruction.

DISCUSSION

During 1964-1965, twelve percent of the Lippincott pupils were retained in Grade I compared to three percent of the i/t/a-Merrill pupils, six percent of the Scott, Foresman pupils and six percent of the Phonics and Word Power pupils. In the second year of the major study, 1965-1966, there were almost eight percent of the Lipp children retained in Grade II compared to almost five percent i/t/a-Merr pupils, almost two percent SF pupils, and almost five percent PWP pupils. In the third year, 1966-1967, there were no retainees at the end of Grade III for Lipp, while there were three percent for i/t/a-Merr, two percent for SF and six percent for PWP.

In the replicative study the retainee percentages in Grade I of 1965-1966 were: 11.3 Lipp, 18.3 i/t/a-Merr, 1.5 SF, and 5.2 PWP. At the end of Grade II of 1966-1967 in the replicative study the retainee percentages were: 3.5 Lipp, 2.1 i/t/a-Merr, 3.2 SF, and 2.1 PWP.

A majority of the retained students attended schools which were located in lower socio-economic areas of New Castle. Their IQ and reading readiness scores, while somewhat lower than the means attained by the entire population, were frequently high enough to suggest that many of the retainees should have succeeded. In an effort to better understand pupil retentions in this study, an additional investigation was conducted in May of 1967 (Appendix F). The high retention in some treatment groups possibly affected the achievement test relative standings in subsequent grades.

Each of the four approaches to teaching beginning reading were used in this study under rather ideal conditions. The in-service education provided the teachers was generally excellent. More than the usual amount of teacher in-service education was provided. Teachers received more supervision than is normally available All of the most recent materials offered by the involved companies were provided. It cannot be assumed that any one of the approaches, without the conditions of this study, would produce the same results.

CONCLUSIONS

By the end of Grade III the Lippincott program produced the best overall results on a standardized silent reading achievement test. The third grade silent achievement test results were as follows in June 1967: (1) Lipp was significantly higher than SF and PWP in Paragraph Meaning, Spelling, Word Meaning and Word Study Skills; and (2) i/t/a-Merr and PWP were significantly higher than SF in Word Study Skills.

By the end of Grade III the i/t/a-Merrill program produced the best overall results on a standardized oral reading achievement test. The third grade oral achievement test results were as follows in April 1967: i/t/a-Merr was significantly higher than both Lipp and SF in oral comprehension.

In Grade III, Lipp and SF pupils read significantly more books than did i/t/a-Merr and PWP pupils. The preceding statement refers to books read other than the regular textbooks.

For the low IQ third in Grade III the only significant difference was in oral comprehension with i/t/a-Merr ahead of Lipp. For the average IQ third in Grade III the only significant difference was in Word Study Skills in favor of Lipp over SF. For the high IQ third in Grade III, Lipp was significantly higher than PWP in Paragraph Meaning and in oral comprehension, i/t/a-Merr was significantly higher than PWP in oral comprehension, while both Lipp and i/t/a-Merr were significantly higher than SF in both oral accuracy and oral comprehension.

In Grade III teachers whose classes were judged to be achieving above their reading expectancy levels were teachers rated as superior in personality and as above average in planning, knowledge, communicative ability, classroom management and attainment of objectives. Teachers of students with above average attitude toward reading were teachers rated above average in all categories.

In the replicative study in Grade II there were no significant differences in oral achievement, while the silent achievement test at the end of the year indicated these significant differences: (1) For the low IQ third, Lipp over i/t/a-Merr in Paragraph Meaning and Word Meaning, while SF was also higher than i/t/a-Merr in Word Meaning; (2) for the average IQ third, i/t/a-Merr over PWP in Paragraph Meaning, Word Study Skills and Spelling, while i/t/a-Merr also was ahead of SF in Word Study Skills; (3) for the average IQ third, Lipp over SF and PWP in Word Study Skills and Lipp over PWP in Spelling; and (4) for the high IQ third, Lipp was ahead of SF in Word Meaning and Spelling and Lipp was also ahead of PWP in Spelling. For the number of books read other than regular textbooks, Lipp read significantly more than SF, while PWP read significantly more than both SF and i/t/a-Merr. There were only slight differences in teacher characteristics of teachers whose classes were above average in achievement and attitude toward reading and teachers whose classes were not above average in achievement or attitude.

IMPLICATIONS

It appears that method and materials as well as teachers can make a difference in the teaching of reading since four of five Lippincott classes, three of five i/t/a-Merrill classes, only two of five Phonics and Word Power classes and only two of the four Scott, Foresman classes had fifty percent or more pupils achieving at least one-half grade above their predicted levels in Word Meaning, Word Study Skills and Paragraph Meaning in June of third grade. Intensive phonic approaches seem to produce significantly better results in Word Study Skills than does an eclectic basal reader. Pupils introduced to reading through the i/t/a-Merrill program are not confused in the area of Spelling. Since by the end of Grade III the i/t/a-Merrill group generally achieved the best results in oral comprehension and the Lippincott group generally achieved the best results in silent reading achievement, it is indicated than an i/t/a-Lippincott program would be worthy of attention and future study.

The generally higher coefficients of correlations in Grade III compared to Grade II between results on the San Diego Attitude Scale and other criterion variables (.56 with Gates Word List) is an indication of more validity for this type of attitudinal instrument in Grade III than in Grade II. Further support to the preceding statement may be found in the low correlation coefficients in Grade I and II in the first two years of this study between the San Diego Attitude Scale and other criterion variables.

SUMMARY

The primary goal was to determine which of four approaches to beginning reading instruction was the most effective. It involved a three-year longitudinal study in which students who were in second grade in 1965-1966 were followed into third grade in 1966-1967. Also, a modified replication was conducted in which first-grade students of 1965-1966 were followed into Grade II in 1966-1967.

The independent treatment variables for both Grades II and III were: (1) A basal reader program using materials published by Scott, Foresman and Company, 1962 edution: (2) a phonic program which utilized correlated filmstrips and published by the J. B. Lippincott Company, 1963 edition; (3) an eclectic, combination program which used the materials of the Scott, Foresman Company (No. 1 above) supplemented with phonic booklets (Phonics and Word Power) published by American Education Publications, Inc.; and (4) a language arts approach using the initial teaching alphabet as a medium, represented by the materials of i/t/a Publications, Inc., 1963 edition. Treatment variable number four (i/t/a) is the only one which has been changed. This was necessary because that program was conceived and designed as one to be used only for the initial teaching of reading. By the time most children completed the first grade, they had transferred to traditional orthography. For this reason, Dr. Albert J. Mazurkiewicz, co-author of the Early-to-Read i/t/a Program and consultant to the first-grade i/t/a classes during the 1964-1965 school year, recommended the Treasury of Literature Series, *published by Charles E. Merrill Books, Inc.*, as appropriate materials to use following transition from i/t/a.

The objectives were:

1. Which of these methods was best for children of different ability levels?
2. What were the teaching characteristics of teachers whose students achieved above their reading expectancy levels?
3. What were the teaching characteristics of teachers whose students maintained high interest in reading?
4. What were the effects of each method on spelling achievement?

The criterion variables were the Stanford Achievement Test, the San Diego Pupil Attitude Inventory, the number of books children read independently, the Gates Word Pronunciation Test, and the Gilmore Oral Reading Test.

Teachers used only those materials and methods recommended by the book companies. Book company consultants provided appropriate in-service education. Teachers were visited frequently to check on adherence to materials and methods. Teachers submitted activity logs to the field director as another method safeguard.

Results were analyzed by analysis of variance and covariance. Pupils were divided into IQ thirds for the analysis of variance and covariance. Coefficients of correlations were computed to determine relationships among variables. Also, Bond and Tinker reading expectancy scores were compared to grade scores for Word Reading, Word Study Skills, and Paragraph Meaning.

At the end of Grade III the Lippincott program produced the best overall results on a standardized silent reading achievement test, while the i/t/a-Merrill produced the best overall results on a standardized oral reading achievement test. In Grade III the Lippincott and Scott, Foresman pupils read significantly more books than did the i/t/a-Merrill and Phonics and Word Power pupils. Four of the five Lippincott classes, three of the five i/t/a-Merrill classes, two of the five Phonics and Word Power classes, and two of the four Scott, Foresman classes had 50 percent or more pupils achieving at least one-half grade above the predicted levels in Word Meaning, Word Study Skills and Paragraph Meaning by June of third grade.

In the replicative study in Grade II there were no significant differences in oral achievement. On the silent achievement test at the end of the year, the Word Meaning scores were significantly higher for Lippincott than for each of the other three programs. The Paragraph Meaning results were significantly higher for Lippincott compared to Scott, Foresman and Phonics and Word Power, while i/t/a-Merrill was significantly higher than Phonics and Word Power. The Word Study Skills scores were significantly higher for Lippincott compared to each of the other three programs, while i/t/a-Merrill was significantly higher than Scott, Foresman and Phonics and Word Power. The Spelling results were significantly higher for Lippincott than for each of the other three programs, while i/t/a-Merrill was significantly higher than Scott, Foresman and Phonics and Word Power. Phonics and Word Power pupils read significantly more books than did i/t/a-Merrill and Scott, Foresman pupils, while the Lippincott pupils also read significantly more books than Scott, Foresman pupils in Grade II. Two of the three Lippincott classes, one of the three i/t/a-Merrill classes, none of the three Scott, Foresman classes, and none of the three Phonics and Word Power classes had 50 percent or more pupils achieving at least one-half grade above their predicted levels for Word Meaning, Word Study Skills and Paragraph Meaning in May of second grade.

The high retention ratios in some Lippincott and i/t/a-Merrill classes possibly affected relative achievement standings in subsequent grades. Intensive phonic approaches seem to produce significantly better results in Word Study Skills than does a typical eclectic basal reader. For Paragraph Meaning the Lippincott program appears to produce significantly better results in Grades II and III than does a typical eclectic basal reader. Pupils introduced to reading through the i/t/a-Merrill program are not confused in the area of Spelling. The Lippincott program appears to be consistently and especially effective in challenging pupils in the high IQ third to high achievement. Since by the end of Grade III the i/t/a-Merrill group generally achieved the best results in oral comprehension and the Lippincott group generally achieved the best results in silent reading achievement, it is indicated that an i/t/a-Lippincott program would be worthy of attention and future study.

REFERENCES

BOOKS

1. Bond, Guy S. and Tinker, Miles A. Reading Difficulties - Their Diagnosis and Correction. New York, New York: Appleton-Century Crofts, Inc. 1957.

2. Jacobs, Leland B. and Turner, Jasper Jo. Treasury of Literature Series. Columbus, Ohio: Charles E. Merrill Books, Inc. 1960.

3. Johnson, Eleanor; Singleton, Carlton; and Wonsavage, Elaine. Phonics and Word Power, Programs 2 and 3. Columbus, Ohio: American Education Publications, Inc. 1964.

4. McCracken, S. Glenn and Walcutt, Charles C. Basic Reading. Philadelphia, Pennsylvania: J. B. Lippincott Company. 1963.

5. Robinson, Helen M.; Monroe, Marion; Artley, Sterl A.; and Greet, Cabell W. The New Basic Readers. Chicago, Illinois: Scott, Foresman and Co. 1961-62.

6. Shelby, Weast, Shankland, and Hodgeman. Handbook of Mathematical Tables. Supplement to Handbook of Chemistry and Physics. Cleveland, Ohio: Chemical Rubber Publishing Company. 1963.

7. Tanyzer, Harold J. and Mazurkiewicz, Albert J. Early to Read i/t/a Program. New York, New York: i/t/a Publications, Inc. 1964.

8. Winer, B. J. Statistical Principles in Experimental Design. New York, New York: McGraw-Hill Book Company. 1962.

ARTICLES

9. Campbell, Donald T. and Stanley, Julian C. "Experimental and Quasi-Experimental Designs for Research on Teaching", Handbook of Research on Teaching. Chicago, Illinois: Rand McNally and Company. 1963.

10. Cook, Desmond L. "The Hawthorne Effect in Educational Research". Phi Delta Kappan, XLIV, 1962.

Articles (Continued)

11. Fay, Leo. "A Look at Two Approaches to the Teaching of Reading". *Changing Concepts of Reading Instruction International Reading Association Conference Proceedings*, VI, Scholastic Magazines, New York, New York. 1961.

12. Gans, Roma. "The Effect of Current Emphases on Beginning Reading Materials and Methods on the Recognition of Individual Differences". *Combining Research Results and Good Practice*. Proceedings 1966, II, Part 2, Eleventh Annual Convention, International Reading Association.

13. Gray, William S. "The Teaching of Reading". *Encyclopedia of Educational Research*. New York, New York: The Macmillan Company. Third Edition, 1960.

14. Harris, Theodore L.; Otto, Wayne; and Barrett, Thomas C. "Summary and Review of Investigations relating to Reading - July 1, 1965 to June 30, 1966". *The Journal of Educational Research*, LX, Number 7, March 1967.

15. Holt, Howard B. "The Educational Uses of Change". *Phi Delta Kappan*, XLVI, December 1964.

16. Kelley, Barbara C. "The Economy Method vs. the Scott, Foresman Method in Teaching Second Grade Reading in the Murphysboro Public Schools". *The Journal of Educational Research*, LI, May - June 1958.

17. Kinsella, John J. "The Role of the State Supervisor in Encouraging Research and Implementing Research Findings". *The Leadership Role of State Supervisors of Mathematics*. Bulletin OE 29032, Washington, D.C.: U.S. Government Printing Office, 1962.

18. Lumsdaine, A. A. "Instruments and Media of Instruction". *Handbook of Research on Teaching*, ed. N. L. Gage, Chicago, Illinois: Rand McNally and Company. 1963.

RESEARCH REPORTS

19. Greenman, Margaret H. "A Six Year Experimental Study of Two Methods of Teaching Reading in the Elementary School", paper presented at a joint meeting of the International Reading Association and the American Research Association, February 17, 1959.

Research Reports (Continued)

20. Harris, Albert J., and Serwer, Blanche L. *Comparison of Reading Approaches in First Grade Teaching with Disadvantaged Children*. The Research Foundation of the City University of New York, New York, New York: USOE Cooperative Research Project No. 2677, 1966.

21. Hayes, Robert B., and Nemeth, Joseph S. *An Attempt to Secure Additional Evidence Concerning Factors Affecting Learning to Read*. Report of Cooperative Research Project No. 2697. U.S. Office of Education, 1964-65.

22. Hayes, Robert B., and Wuest, Richard C. *Factors Affecting Learning to Read*. Report of Cooperative Research Project No. 5-0572, U.S. Office of Education. December 1966.

APPENDIX A

HAYES TEACHER RATING SCALE

Observer's Name____________________ School____________________

Teacher Being Observed________________ Date of Observation______________

Reading Method_____________________ Observation No._________________

Directions: Your observation of each lesson should determine whether it is above average, average, or below average. In left-hand column, enter (M) for adequate, (+ or -) for deviations above and below; for example, for any one of the four sub-points under category A. PLANNING, your observation can be M, M+, or M-. Make appropriate comments on the right if a statement under each of the six categories does not appear on the form. After the lesson, circle the appropriate number under Rating of each of the six categories to indicate your overall rating of that category (a rating of 2 represents an average performance).

After you have filled out this Performance Method sheet, you are to send it to the School Psychologist, Mr. DeCaprio, at the Administration Building, and mark it Confidential Material.

New Castle Reading Study

APPENDIX A (CONTINUED)

HAYES TEACHER RATING SCALE

	RATING	COMMENTS	
		STRENGTHS	WEAKNESSES
A. PLANNING	0 1 2 3 4		
____1. Motivated students			
____2. Adequate student performance time			
____3. Suited amount of material to time			
____4. Established closure			
B. KNOWLEDGE	0 1 2 3 4		
____1. Explanations were correct			
____2. Answered student questions adequately			
____3. Student supervision indicated thorough knowledge			
C. COMMUNICATION TO STUDENTS	0 1 2 3 4		
____1. Emphasized main points			
____2. Provided good examples			
____3. Used standard English			
____4. Varied speech patterns (rate and volume)			
____5. Effectively used training aids			
D. CLASSROOM MANAGEMENT	0 1 2 3 4		
____1. Effectively arranged physical facilities			
____2. Channeled student activity toward objectives			
____3. Used supervision for additional teaching			
____4. Included an evaluation step			
E. TEACHING PERSONALITY	0 1 2 3 4		
____1. Interested in students			
____2. Interested in subject			
____3. Attitude which fostered cooperation			
____4. Confident and at ease			
F. ATTAINMENT OF OBJECTIVES	0 1 2 3 4		

New Castle Reading Study

APPENDIX A (CONTINUED)

HAYES TEACHER RATING SCALE
RATING RANGE CATEGORIES

Rating Range	Comments
Planning	
3.0-4.0	Superior planning of all aspects
2.0-2.9	Above average planning
1.0-1.9	About average in planning lessons
0 - .9	Definitely below average in planning lessons
Knowledge	
3.0-4.0	Superior knowledge of subject
2.0-2.9	Above average knowledge of subject
1.0-1.9	About average knowledge of subject
0 - .9	Definitely below average knowledge of subject
Communication	
3.0-4.0	Superior communication of ideas
2.0-2.9	Above average communication of ideas
1.0-1.9	About average communication of ideas
0 - .9	Definitely below average
Management	
3.0-4.0	Superior guidance, supervision and evaluation of students toward lesson objectives
2.0-2.9	Above average supervision
1.0-1.9	About average supervision
0 - .9	Poor
Personality	
3.0-4.0	Superior attitude which completely gained pupil cooperation
2.0-2.9	Above average personality
1.0-1.9	About average
0 - .9	Poor
Objectives	
3.0-4.0	Superior teacher-pupil achievement
2.0-2.9	Above average teacher-pupil achievement
1.0-1.9	About average teacher-pupil achievement
0 - .9	Definitely below average teacher-pupil achievement

APPENDIX B

GRADE III TEACHER LOG

Teacher____________________ School ____________ Treatment ____________ Date __________________

	Whole Class	Group I	Group II	Group III	Total
I. Reading Instructional Time					
II. Skills Taught					
III. Materials					
A. Basal Readers					
1. Title					
2. Level					
3. Pages					
B. Workbooks					
1. Level					
2. Pages					
C. Filmstrips					
1. Number					
2. Frames					
D. Other					
1. Title					
2. Level					
3. Pages					
IV. Number of students					

Visitors Present
A. Name ____________________
B. Length of visit ____________________
C. Comments ____________________

APPENDIX C

GRADE II TEACHER LOG

Date ________________

Teacher ________________ School ____________ Treatment ________

I. Materials
  A. Books (Include textbooks, workbooks, practice pads, etc.)
    1. Entire class Number of students__________
      Titles Level Pages
      a.
      b.
      c.
      d.
      e.
    2. Group I Number of students __________
      Titles Level Pages
      a.
      b.
      c.
      d.
      e.
    3. Group II Number of students __________
      Titles Level Pages
      a.
      b.
      c.
      d.
      e.
    4. Group III Number of students __________
      Titles Level Pages
      a.
      b.
      c.
      d.
      e.
  B. Motion pictures, slides, filmstrips, etc.
    1. Entire class
      Titles Level (if available)
      a.
      b.
      c.
      d.
      e.

APPENDIX C (CONTINUED)

2. Group I

Titles	Level (if available)

a.
b.
c.
d.
e.

3. Group II

Titles	Level (if available)

a.
b.
c.
d.
e.

4. Group III

Titles	Level (if available)

a.
b.
c.
d.
e.

II. Visitors Present

A. Name________________________________

B. Length of Visit ____________________

III. Comments

APPENDIX D

TABLE 1
JANUARY 1967 - GRADE II - WORD MEANING
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	21.46 (3.1)	4.19**	5.72**	7.74**
i/t/a-Merr	17.27 (2.7)		1.53	3.55*
SF	15.74 (2.7)			2.02
PWP	13.72 (2.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 students was 17.50 (2.8) with a standard deviation of 7.03.

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 2
JANUARY 1967 - GRADE II - WORD MEANING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	1407.507	469.169	14.089**
IQ Levels	2	122.141	61.070	1.834
Interaction	6	790.887	131.814	3.958**
Error	166	5527.405	33.297	
Total	177	7847.942		

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 3
JANUARY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

	Means[1]	Differences i/t/a-Merr	 SF	 PWP
Lipp	33.09 (3.0)	5.14	7.69**	11.02**
i/t/a-Merr	27.95 (2.7)		2.55	5.88*
SF	25.40 (2.5)			3.33
PWP	22.07 (2.4)			

[1] Raw score means followed by grade equivalent means. The grand mean was 27.13 (2.6) for all 180 pupils with a standard deviation of 11.78.

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 4
JANUARY 1967 - GRADE II - PARAGRAPH MEANING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	2817.764	939.254	9.266**
IQ Levels	2	372.315	186.157	1.836
Interaction	6	1449.321	241.553	2.383
Error	166	16825.985	101.361	
Total	177	21465.385		

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 5
JANUARY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	45.76 (4.0)	3.12	12.94**	14.51**
i/t/a-Merr	42.64 (3.6)		9.82**	11.39**
PWP	32.82 (2.7)			1.57
SF	31.25 (2.5)			

[1] Raw score means followed by grade equivalent means. The grand mean was 38.12 (3.1) for all 180 pupils with a standard deviation of 12.24.

** Significance exceeds at .01 level.

TABLE 6
JANUARY 1967 - GRADE II - WORD STUDY SKILLS
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	6440.542	2146.847	24.358**
IQ Levels	2	354.191	177.095	2.009
Interaction	6	1307.675	217.945	2.472*
Error	166	14630.649	88.136	
Total	177	22733.057		

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 7
JANUARY 1967 - GRADE II - SPELLING
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	18.69 (3.6)	2.79	9.00**	10.05**
i/t/a-Merr	15.90 (3.3)		6.21	7.26*
SF	9.69 (2.6)			1.05
PWP	8.64 (2.5)			

[1] Raw score means followed by grade equivalent means. The grand mean was 13.23 (3.0) for all 180 pupils with a standard deviation of 8.43.

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 8
JANUARY 1967 - GRADE II - SPELLING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	3043.875	1014.625	7.409**
IQ Levels	2	259.968	129.984	0.949
Interaction	6	821.662	136.943	2.812*
Error	166	8083.461	48.695	
Total	177	12208.967		

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 9
JANUARY 1967 - GRADE II - WORD MEANING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	23.53 (3.6)	2.00	6.40*	7.73**
i/t/a-Merr	21.53 (3.3)		4.40	5.73*
SF	17.13 (2.7)			1.33
PWP	15.80 (2.7)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the high IQ third was 19.50 (3.0).

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

TABLE 10
JANUARY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	38.73 (3.3)	4.60	11.46*	11.73*
i/t/a-Merr	34.13 (3.0)		6.86	7.13
PWP	27.27 (2.6)			.27
SF	27.00 (2.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the high IQ third was 31.78 (2.9).

* Significance exceeds at .05 level.

TABLE 11
JANUARY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	50.27 (4.8)	4.00	11.74**	13.74**
i/t/a-Merr	46.27 (4.0)		7.74	9.74*
SF	38.53 (3.2)			2.00
PWP	36.53 (3.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the high IQ third was 42.90 (3.6).

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 12
JANUARY 1967 - GRADE II - SPELLING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	21.67 (3.9)	2.94	10.34**	11.27**
i/t/a-Merr	18.73 (3.6)		7.40*	8.33**
SF	11.33 (2.8)			1.74
PWP	10.40 (2.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the high IQ third was 15.53 (3.3).

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 13
JANUARY 1967 - GRADE II - WORD MEANING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	SF	PWP
Lipp	23.53 (3.6)	2.13	8.06**	9.97**
i/t/a-Merr	21.40 (3.1)		5.93*	7.80**
SF	15.47 (2.6)			1.87
PWP	13.60 (2.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the average IQ third was 17.98 (2.8).
** Significance exceeds at .01 level.
* Significance exceeds at .05 level.

TABLE 14
JANUARY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	SF	PWP
i/t/a-Merr	34.20 (3.0)	2.27	7.80	12.13**
Lipp	31.93 (2.9)		5.53	9.86*
SF	26.40 (2.6)			4.33
PWP	22.07 (2.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the average IQ third was 28.65 (2.7).
** Significance exceeds at .01 level.
* Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 15
JANUARY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	SF	PWP
i/t/a-Merr	48.27 (4.5)	1.14	16.34**	16.74**
Lipp	47.13 (4.2)		15.20**	15.60**
SF	31.93 (2.6)			.40
PWP	31.53 (2.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the average IQ third was 39.72 (3.3).

** Significance exceeds at .01 level.

TABLE 16
JANUARY 1967 - GRADE II - SPELLING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	SF	PWP
i/t/a-Merr	20.73 (3.8)	.93	10.60**	12.93**
Lipp	19.80 (3.7)		9.67**	12.00**
SF	10.13 (2.6)			2.33
PWP	7.80 (2.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the average IQ third was 14.62 (3.2).

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 17
JANUARY 1967 - GRADE II - WORD MEANING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	PWP	i/t/a-Merr
Lipp	18.40 (2.8)	4.40	5.40	9.13
SF	14.00 (2.5)		1.00	4.73
PWP	13.00 (2.3)			3.73
i/t/a-Merr	9.27 (1.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the low IQ third was 13.67 (2.5).

TABLE 18
JANUARY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	PWP	i/t/a-Merr
Lipp	26.80 (2.6)	5.47	7.53	10.40*
SF	21.33 (2.3)		2.06	4.93
PWP	19.27 (2.1)			2.87
i/t/a-Merr	16.40 (1.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the low IQ third was 20.95 (2.3).

* Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 19
JANUARY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences i/t/a-Merr	 PWP	 SF
Lipp	38.33 (3.1)	7.60	8.40	10.40*
i/t/a-Merr	30.73 (2.5)		.80	2.80
PWP	29.93 (2.4)			2.00
SF	27.93 (2.3)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the low IQ level was 31.73 (2.6).

* Significance exceeds at .05 level.

TABLE 20
JANUARY 1967 - GRADE II - SPELLING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences SF	 i/t/a-Merr	 PWP
Lipp	14.13 (3.1)	5.87	6.26	6.26
SF	8.26 (2.4)		.39	.39
i/t/a-Merr	7.87 (2.4)			.00
PWP	7.87 (2.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 pupils in the low IQ third was 9.53 (2.6).

TABLE 21
JANUARY 1967 - GRADE II - WORD MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	1242.151	414.050	12.016**
IQ Levels	2	1099.233	549.616	15.951**
Interaction	6	728.368	121.394	3.523**
Error	168	5788.799	34.457	
Total	179	8858.551		

** Significance exceeds at .01 level.

TABLE 22
JANUARY 1967 - GRADE II - PARAGRAPH MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	2387.660	795.886	7.565**
IQ Levels	2	3729.380	1864.690	17.521**
Interaction	6	1287.290	214.548	2.039
Error	168	17677.730	105.224	
Total	179	25082.060		

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 23
JANUARY 1967 - GRADE II - WORD STUDY SKILLS
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	5490.720	1830.240	19.238**
IQ Levels	2	3971.230	1985.615	20.870**
Interaction	6	1155.810	192.635	2.025
Error	168	15982.780	95.135	
Total	179	26600.540		

** Significance exceeds at .01 level.

TABLE 24
JANUARY 1967 - GRADE II - SPELLING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	2596.549	856.441	18.068**
IQ Levels	2	1140.308	561.323	11.842**
Interaction	6	678.145	113.024	2.384*
Error	168	7678.952	47.401	
Total	179	12093.954		

** Significance exceeds at .01 level.
 * Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 25
JANUARY 1967 - GRADE II - PARAGRAPH MEANING CELL DATA FOR UNWEIGHTED MEANS ANALYSIS

			IQ LEVELS		
			b_1 (High)	b_2 (Average)	b_3 (Low)
TREATMENTS	a_1 (SF)	n_{ij}	30	15	15
		SX	898	396	320
		SX^2	31278	11490	8280
		SS_{ij}	4397.86667	1035.6	1453.33334
	a_2 (Lipp)	n_{ij}	16	19	21
		SX	631	613	542
		SX^2	26599	21201	16470
		SS_{ij}	1713.9375	1423.68422	2481.2381
	a_3 (PWP)	n_{ij}	17	17	15
		SX	485	363	289
		SX^2	16195	9135	6481
		SS_{ij}	2358.2353	1383.88236	912.93334
	a_4 (i/t/a-Merr)	n_{ij}	15	15	16
		SX	512	513	263
		SX^2	18590	19167	5857
		SS_{ij}	1113.73334	1622.4	1533.9375

APPENDIX D (CONTINUED)

TABLE 26
JANUARY 1967 - GRADE II - PARAGRAPH MEANING
CELL MEANS

	IQ LEVELS			
TREATMENTS	b_1 (High)	b_2 (Average)	b_3 (Low)	Total
a_1 (SF)	29.93	26.40	21.33	77.66
a_2 (Lipp)	39.44	32.26	25.81	97.51
a_3 (PWP)	28.53	21.35	19.27	69.15
a_4 (i/t/a-Merr)	34.13	34.20	16.44	84.77
	132.03	114.21	82.85	329.09

TABLE 27
SUMMARY OF ANALYSIS OF VARIANCE

Source of Variation	SS	df	MS	F
A (Method)	2435.571760	3	811.85725	7.53867**
B (IQ Level)	5244.071419	2	2622.03571	24.34746**
AB	1303.372894	6	217.22882	2.01712
Within cell	21430.78167	199	107.69237	

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 28
JANUARY 1967 - GRADE II - PARAGRAPH MEANING

ALL 211 PUPILS

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	31.89 (2.9)	3.89	4.99*	8.69**
i/t/a-Merr	28.00 (2.7)		1.10	4.80
SF	26.90 (2.6)			3.70
PWP	23.20 (2.4)			

HIGH IQ

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	39.44 (3.3)	5.31	9.51*	10.91*
i/t/a-Merr	34.13 (3.0)		4.20	5.60
SF	29.93 (2.8)			1.40
PWP	28.53 (2.7)			

AVERAGE IQ

		Differences		
	Means[1]	Lipp	SF	PWP
i/t/a-Merr	34.20 (3.0)	1.94	7.80	12.85**
Lipp	32.26 (2.9)		5.86	10.91*
SF	26.40 (2.6)			5.05
PWP	21.35 (2.3)			

LOW IQ

		Differences		
	Means[1]	SF	PWP	i/t/a-Merr
Lipp	25.81 (2.6)	4.48	6.54	9.37*
SF	21.33 (2.3)		2.06	4.89
PWP	19.27 (2.1)			2.83
i/t/a-Merr	16.44 (1.9)			

[1] Raw score means followed by grade equivalent means.

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 29
MAY 1967 - GRADE II - WORD MEANING
COMPARISON OF ADJUSTED MEANS

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	22.49 (3.3)	2.62*	3.59**	3.79**
i/t/a-Merr	19.87 (3.0)		.97	1.17
PWP	18.90 (2.9)			.20
SF	18.70 (2.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 19.99 (3.0) with a standard deviation of 6.94.

* Significant at .05 level.

** Significant at .01 level.

TABLE 30
MAY 1967 - GRADE II - WORD MEANING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	398.755	132.918	5.104**
IQ Levels	2	89.701	44.850	1.722
Interaction	6	447.394	74.565	2.864*
Error	167	4348.593	26.039	
Total	178	5284.443		

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 31
MAY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF ADJUSTED MEANS

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	39.42 (3.3)	3.17	6.01**	7.08**
i/t/a-Merr	36.25 (3.1)		2.84	3.91*
SF	33.41 (3.0)			1.07
PWP	32.34 (2.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 35.36 (3.1) with a standard deviation of 10.00.

** Significance exceeds at .01 level.
 * Significance exceeds at .05 level.

TABLE 32
MAY 1967 - GRADE II - PARAGRAPH MEANING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	1326.552	442.184	6.492**
IQ Levels	2	319.312	159.656	2.344
Interaction	6	1130.947	188.491	2.767*
Error	167	11374.201	68.108	
Total	178	14151.012		

** Significance exceeds at .01 level.
 * Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 33
MAY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF ADJUSTED MEANS

	Means[1]	Differences i/t/a-Merr	SF	PWP
Lipp	51.37 (5.0)	4.52*	11.54**	12.42**
i/t/a-Merr	46.85 (4.2)		7.02**	7.90**
SF	39.83 (3.3)			.88
PWP	38.95 (3.2)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 44.25 (3.7) with a standard deviation of 8.99.

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

TABLE 34
MAY 1967 - GRADE II - WORD STUDY SKILLS
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	4566.655	1522.218	13.185**
IQ Levels	2	181.914	90.957	0.788
Interaction	6	1247.867	207.978	1.801
Error	167	19280.489	115.452	
Total	178	25276.925		

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 35
MAY 1967 - GRADE II - SPELLING
COMPARISON OF ADJUSTED MEANS

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	22.14 (3.9)	3.30*	6.09**	6.27**
i/t/a-Merr	18.84 (3.6)		2.79*	2.97*
SF	16.05 (3.3)			.18
PWP	15.87 (3.3)			

[1] Raw score means followed by grade equivalent means. The grand mean was 18.22 (3.5) with a standard deviation of 7.30.

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

TABLE 36
MAY 1967 - GRADE II - SPELLING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatment	3	1134.396	378.132	8.947**
IQ Levels	2	77.573	38.786	0.918
Interaction	6	479.414	79.902	1.891
Error	167	7058.319	42.265	
Total	178	8749.702		

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 37
MAY 1967 - GRADE II - LANGUAGE
COMPARISON OF ADJUSTED MEANS

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	45.60 (3.6)	2.90	5.74**	7.22**
i/t/a-Merr	42.70 (3.3)		2.84	4.32**
SF	39.86 (3.1)			1.48
PWP	38.38 (2.9)			

[1] Raw score means followed by grade equivalent means. The grand mean was 41.63 (3.2) with a standard deviation of 9.52.

** Significance exceeds .01 level.

TABLE 38
MAY 1967 - GRADE II - LANGUAGE
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	1345.490	448.497	7.22**
IQ Levels	2	237.237	118.619	1.909
Interaction	6	786.501	131.084	2.110
Error	167	10374.578	62.123	
Total	178	12743.806		

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 39
MAY 1967 - GRADE II - WORD MEANING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	24.73 (3.7)	1.93	3.06	5.06*
i/t/a-Merr	22.80 (3.5)		1.13	3.13
PWP	21.67 (3.3)			2.00
SF	19.67 (3.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ level pupils was 22.22 (3.3).

* Significance exceeds at .05 level.

TABLE 40
MAY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	SF	PWP
Lipp	43.47 (3.7)	1.67	5.94	7.40
i/t/a-Merr	41.80 (3.6)		4.27	1.46
SF	37.53 (3.2)			1.46
PWP	36.07 (3.1)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 39.72 (3.4).

APPENDIX D (CONTINUED)

TABLE 41
MAY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	54.53 (5.8)	4.60	6.06	9.60
i/t/a-Merr	49.93 (4.8)		1.46	5.00
SF	48.47 (4.5)			3.54
PWP	44.93 (3.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 49.47 (4.7).

TABLE 42
MAY 1967 - GRADE II - SPELLING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	25.07 (4.4)	4.34	6.47*	7.80**
i/t/a-Merr	20.73 (3.8)		2.13	3.46
SF	18.60 (3.6)			1.33
PWP	17.27 (3.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 20.42 (3.7)

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 43
MAY 1967 - GRADE II - LANGUAGE
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences i/t/a-Merr	SF	PWP
Lipp	51.47 (4.2)	4.67	5.60	10.67**
i/t/a-Merr	46.80 (3.7)		0.93	6.00
SF	45.87 (3.6)			5.07
PWP	40.80 (3.1)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 46.23 (3.6).

** Significance exceeds at .01 level.

TABLE 44
MAY 1967 - GRADE II - WORD MEANING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences Lipp	SF	PWP
i/t/a-Merr	22.87 (3.5)	0.54	3.47	4.47
Lipp	22.33 (3.3)		2.93	3.93
SF	19.40 (2.9)			1.00
PWP	18.40 (2.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 20.75 (3.1).

APPENDIX D (CONTINUED)

TABLE 45
MAY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences Lipp	 SF	 PWP
i/t/a-Merr	40.73 (3.5)	0.93	5.46	8.93*
Lipp	39.80 (3.4)		4.53	8.00
SF	35.27 (3.1)			3.47
PWP	31.80 (2.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 36.90 (3.2).

* Significance exceeds at .05 level.

TABLE 46
MAY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences i/t/a-Merr	 SF	 PWP
Lipp	54.93 (5.8)	3.33	15.60**	19.40**
i/t/a-Merr	51.60 (5.2)		12.27*	16.07**
SF	39.33 (3.2)			3.80
PWP	35.53 (2.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 49.47 (4.7).

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 47
MAY 1967 - GRADE II - SPELLING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	SF	PWP
Lipp	23.07 (4.0)	1.34	6.07	8.94**
i/t/a-Merr	21.73 (3.9)		4.73	7.60**
SF	17.00 (3.4)			2.87
PWP	14.13 (3.1)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 18.98 (3.6).

** Significance exceeds at .01 level.

TABLE 48
MAY 1967 - GRADE II - LANGUAGE
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	PWP	SF
i/t/a-Merr	46.73 (3.7)	1.40	6.80	7.06
Lipp	45.33 (3.5)		5.40	5.66
PWP	39.93 (3.1)			0.26
SF	39.67 (3.1)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 42.92 (3.3).

APPENDIX D (CONTINUED)

TABLE 49
MAY 1967 - GRADE II - WORD MEANING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences SF	PWP	i/t/a-Merr
Lipp	20.00 (3.0)	1.80	3.53	6.67**
SF	18.20 (2.8)		1.73	4.87*
PWP	16.47 (2.7)			3.14
i/t/a-Merr	13.33 (2.3)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 17.00 (2.7).

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 50
MAY 1967 - GRADE II - PARAGRAPH MEANING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences SF	PWP	i/t/a-Merr
Lipp	33.67 (3.0)	2.34	5.07	9.47*
SF	31.33 (2.9)		2.73	7.13
PWP	28.60 (2.7)			4.40
i/t/a-Merr	24.20 (2.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 29.45 (2.7).

* Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 51
MAY 1967 - GRADE II - WORD STUDY SKILLS
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	43.40 (3.6)	6.27	7.53	8.07
i/t/a-Merr	37.13 (3.0)		1.26	1.80
PWP	35.87 (2.9)			0.54
SF	35.33 (2.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 37.93 (3.1).

TABLE 52
MAY 1967 - GRADE II - SPELLING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	SF	i/t/a-Merr
Lipp	17.80 (3.5)	1.80	3.87	4.47
PWP	16.00 (3.3)		2.07	2.67
SF	13.93 (3.1)			0.60
i/t/a-Merr	13.33 (3.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 15.27 (3.2).

APPENDIX D (CONTINUED)

TABLE 53
MAY 1967 - GRADE II - LANGUAGE
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	PWP	i/t/a-Merr
Lipp	39.00 (3.0)	2.07	5.00	5.93
SF	36.93 (2.8)		2.93	3.86
PWP	34.00 (2.6)			0.93
i/t/a-Merr	33.07 (2.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 IQ pupils was 35.75 (2.8).

TABLE 54
FEBRUARY 1967 - GRADE II - BOOKS READ
ANALYSIS OF MEANS (ADJUSTED BY COVARIANCE)

		Differences		
	Means[1]	Lipp	i/t/a-Merr	SF
PWP	12.19	2.60	6.02**	7.92**
Lipp	9.59		3.42	5.32**
i/t/a-Merr	6.17			1.90
SF	4.27			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 8.06 with a standard deviation of 7.35.

** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 55
FEBRUARY 1967 - GRADE II - BOOKS READ
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	1495.403	498.468	11.394**
IQ Levels	2	6.600	3.300	0.075
Interaction	6	674.806	112.467	2.571*
Error	166	7262.460	43.750	
Total	177	9439.2698		

** Significance exceeds at .01 level.
* Significance exceeds at .05 level.

TABLE 56
FEBRUARY 1967 - GRADE II - BOOKS READ
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	i/t/a-Merr	SF
PWP	17.27	6.00	10.67**	12.07**
Lipp	11.27		4.67	6.07
i/t/a-Merr	6.60			1.40
SF	5.20			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 10.08.
** Significance exceeds at .01 level.

APPENDIX D (CONTINUED)

TABLE 57
FEBRUARY 1967 - GRADE II - BOOKS READ
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	i/t/a-Merr	SF
PWP	9.00	.27	1.73	2.60
Lipp	8.73		1.46	2.33
i/t/a-Merr	7.27			.87
SF	6.40			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 7.85.

TABLE 58
FEBRUARY 1967 - GRADE II - BOOKS READ
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		PWP	SF	i/t/a-Merr
Lipp	8.93	.73	3.13	6.93*
PWP	8.20		2.40	6.20
SF	5.80			3.80
i/t/a-Merr	2.00			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 6.23.

* Significance exceeds at .05 level.

APPENDIX D (CONTINUED)

TABLE 59
MAY 1967 - GRADE II - SILENT READING ACHIEVEMENT AND READING EXPECTANCY BY CLASSES[1]

Program	Teacher	No. of Pupils at Least 1/2 Grade Above[2]	Total Reading Study Students	50% or More Achieved at Least 1/2 Grade Above Predicted Score
SF	A	3	16	
	B	7	22	
	C	2	24	
Lipp	D	4	14	
	E	10	18	X
	F	14	24	X
SF + PWP	G	6	17	
	H	3	16	
	I	0	14	
i/t/a-Merr	J	13	22	X
	K	5	11	
	L	2	15	
Totals		69	213	3

[1] Comparisons made between predicted achievement on Bond and Tinker Reading Grade Scores and actual achievement on Stanford Achievement Test, Primary II, Form X.

[2] Number includes those who were one-half grade level above the Bond and Tinker Reading Grade Score on at least two of three reading subtests (Word Meaning, Paragraph Meaning, and Word Study Skills - Stanford Achievement Test, Primary II, Form X).

APPENDIX D (CONTINUED)

TABLE 60

TEACHING EFFECTIVENESS[1] AND PUPIL ACHIEVEMENT IN GRADE II

Program	Teacher[2]	Plan	Know	Com	Manage	Pers	Obj	Totals[3]
Lipp	E	2.50	2.75	2.50	2.42	2.92	2.42	15.51
Lipp	F	3.00	3.00	3.17	2.83	3.75	3.17	18.92
i/t/a-Merr	J	1.67	1.67	1.33	1.75	1.67	1.58	9.67
Mean[4]		2.39	2.47	2.33	2.33	2.78	2.39	14.70
Mean[5]		2.40	2.40	2.42	2.38	2.69	2.42	14.71
Total Mean[6]		2.39	2.42	2.37	2.37	2.71	2.41	14.67

[1] Defined by the Hayes Teacher Rating Scale on a 0-4 scale where 2 is average.

[2] Identified in Table 59 and for whom the majority of pupils achieved one-half level above expectancy levels.

[3] Based upon 0-24 point scale.

[4] Scores attained by teachers who had 50% or more reading study students at least one-half grade level above expectancy levels.

[5] Scores attained by teachers who did not have 50% or more reading study students at least one-half grade level above expectancy levels.

[6] Scores attained by all twelve reading study teachers in Grade II.

APPENDIX D (CONTINUED)

TABLE 61
PUPIL ATTITUDE[1] IN GRADE II

Program	Teacher[2]	N. Positive Attitude[3]	Total Reading Study Students	50% or More Positive Attitude[4]
SF	A	12	16	X
	B	14	22	X
	C	12	24	X
Lipp	D	11	14	X
	E	12	18	X
	F	19	24	X
SF + PWP	G	8	17	
	H	12	16	X
	I	8	14	X
i/t/a-Merr	J	8	22	
	K	11	11	X
	L	4	15	
Totals		131	213	9

[1] Measured by San Diego County Inventory of Reading Attitude.

[2] Letter designations are for the same teachers as in Tables 59 and 60.

[3] Number of reading study students in each classroom with stanine score of six or better (raw score 19+) on San Diego County Inventory of Reading Attitude.

[4] Classes where 50% or more of students had above average attitudes toward reading.

APPENDIX D (CONTINUED)

TABLE 62

TEACHING EFFECTIVENESS[1] AND PUPIL ATTITUDE[2] IN GRADE II

Program	Teacher[3]	Plan	Know	Com	Manage	Pers	Obj	Total
SF	A	3.31	3.00	3.31	3.08	3.92	3.08	19.70
	B	1.92	2.08	2.25	2.33	2.33	2.25	13.16
	C	2.92	2.92	2.83	2.92	3.25	2.92	17.76
Lipp	D	1.50	2.00	1.50	1.75	1.83	1.75	10.33
	E	2.50	2.75	2.50	2.42	2.92	2.42	15.51
	F	3.00	3.00	3.17	2.83	3.75	3.17	18.92
SF + PWP	H	2.00	2.08	2.33	1.92	1.58	1.92	11.83
	I	2.83	2.25	2.67	2.25	3.00	2.58	15.58
i/t/a-Merr	K	3.33	3.25	3.00	3.17	3.83	3.33	19.91
Mean[4]		2.59	2.59	2.62	2.52	2.93	2.60	15.85
Mean[5]		1.63	1.72	1.48	1.75	1.85	1.66	10.09
Total Mean[6]		2.39	2.42	2.37	2.37	2.71	2.41	14.67

[1] Teaching Effectiveness as measured by Hayes Teacher Rating Scale.

[2] Pupil Attitude as measured by San Diego County Inventory of Reading Attitude.

[3] Letter designations are for the same teachers in Tables 59, 60, and 61. These are teachers in whose rooms 50% or more of the reading study students scored above average (stanine six, raw score 19+) on the San Diego County Inventory of Reading Attitudes.

[4] Scores attained by teachers whose classes qualified as having a better than average attitude toward reading.

[5] Scores attained by teachers whose classes did not qualify as having a better than average attitude toward reading.

[6] Scores attained by all twelve reading study teachers in Grade II.

APPENDIX D (CONTINUED)

TABLE 63
1966-67 - GRADE II - CORRELATION MATRIX B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. January Word Meaning																			
2. January Paragraph Meaning	88																		
3. January Word Study Skills	73	72																	
4. January Spelling	77	77	75																
5. Books Read	18	23	20	17															
6. IQ	35	37	40	26	24														
7. Teacher Effectiveness Rating	05	03	12	09	17	18													
8. May Word Meaning	80	76	64	67	30	35	07												
9. May Paragraph Meaning	81	81	65	68	25	46	06	84											
10. May Word Study Skills	71	70	78	74	24	38	11	69	71										
11. May Language	67	70	67	65	18	46	19	63	69	67									
12. May Spelling	75	76	65	83	22	28	14	67	70	76	68								
13. May Science and Social Studies	42	44	48	37	24	45	13	51	43	40	45	33							
14. May Arithmetic Computation	45	49	47	43	34	40	32	47	52	54	49	52	32						
15. May Arithmetic Concepts	58	62	60	50	29	56	21	61	65	60	67	59	54	56					
16. San Diego Attitude Scale	12	12	02	13	17	01	28	13	12	13	07	16	00	08	-03				
17. Gilmore Accuracy	67	64	49	58	23	27	15	73	74	61	58	66	17	50	46	-02			
18. Gilmore Comprehension	70	70	55	57	11	35	31	73	65	60	56	63	38	55	54	09	74		
19. Gilmore Rate	43	44	28	33	31	16	08	55	59	37	40	46	-09	34	22	03	57	44	
20. Gates Word List	77	79	73	74	21	25	17	81	77	78	72	85	37	64	53	14	76	77	56

APPENDIX E

TABLE 64
JANUARY 1967 - GRADE III - WORD READING
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	27.82 (4.2)	0.82	2.16	2.73*
i/t/a-Merr	27.00 (4.0)		1.34	1.91
SF	25.66 (3.8)			0.57
PWP	25.09 (3.7)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 26.39 (3.8) with a standard deviation of 5.46.

* Significance exceeds at .05 level.

TABLE 65
JANUARY 1967 - GRADE III - WORD READING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	207.674	69.225	3.146*
IQ Levels	2	21.040	10.520	0.478
Interaction	6	78.095	13.016	0.592
Error	166	3652.671	22.004	
Total	177	3959.480		

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 66
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	44.04 (3.8)	0.19	1.30	4.23
i/t/a-Merr	43.85 (3.8)		1.11	4.04
PWP	42.74 (3.7)			2.93
SF	39.81 (3.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 42.61 (3.7) with a standard deviation of 9.74.

TABLE 67
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	514.619	171.540	2.321
IQ Levels	2	177.349	88.675	1.195
Interaction	6	79.904	13.317	0.180
Error	166	12263.714	73.878	
Total	177	13035.586		

APPENDIX E (CONTINUED)

TABLE 68
JANUARY 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	52.38 (5.2)	1.99	6.43*	9.88**
i/t/a-Merr	50.39 (4.8)		4.44	7.89**
PWP	45.95 (4.0)			3.45
SF	42.50 (3.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 47.79 (4.5) with a standard deviation of 12.68.

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

TABLE 69
JANUARY 1967 - GRADE III - WORD STUDY SKILLS
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	2644.028	881.343	7.343**
IQ Levels	2	221.484	110.742	0.923
Interaction	6	126.768	21.128	0.176
Error	166	19804.136	120.025	
Total	177	22796.416		

** Significance exceeds at .01 level.

TABLE 70
JANUARY 1967 - GRADE III - SPELLING
COMPARISON OF MEANS (ADJUSTED BY COVARIANCE)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	24.83 (4.4)	1.39	3.84*	6.06**
Lipp	23.44 (4.0)		2.45	4.67**
PWP	20.99 (3.8)			2.22
SF	18.77 (3.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 22.00 (3.9) with a standard deviation of 6.59.

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

TABLE 71
JANUARY 1967 - GRADE III - SPELLING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	946.124	315.375	8.599**
IQ Levels	2	80.168	40.084	1.093
Interaction	6	32.329	5.388	0.147
Error	166	6124.791	36.675	
Total	177	7183.412		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 72
JANUARY 1967 - GRADE III - WORD READING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences SF	 i/t/a-Merr	 PWP
Lipp	29.93 (4.7)	0.46	1.26	3.53
SF	29.47 (4.4)		0.80	3.07
i/t/a-Merr	28.67 (4.4)			2.27
PWP	26.40 (3.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 28.61 (4.4).

TABLE 73
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences i/t/a-Merr	 PWP	 SF
Lipp	47.87 (4.3)	2.20	3.34	3.47
i/t/a-Merr	45.67 (4.0)		1.14	1.27
PWP	44.53 (3.9)			0.13
SF	44.40 (3.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 45.61 (4.0).

APPENDIX E (CONTINUED)

TABLE 74
JANUARY 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	55.86 (6.0)	2.60	7.26	7.99
i/t/a-Merr	53.26 (5.4)		4.66	5.39
PWP	48.60 (4.7)			0.73
SF	47.87 (4.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 51.69 (5.2).

TABLE 75
JANUARY 1967 - GRADE III - SPELLING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	25.80 (4.6)	1.20	3.13	5.40
Lipp	24.60 (4.4)		1.93	4.20
PWP	22.67 (4.0)			2.27
SF	20.40 (3.7)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 23.24 (4.0).

TABLE 76
JANUARY 1967 - GRADE III - WORD READING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	28.40 (4.2)	0.47	2.14	2.40
i/t/a-Merr	27.93 (4.2)		1.67	1.93
PWP	26.26 (3.8)			0.26
SF	26.00 (3.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ cases was 27.15 (4.0).

TABLE 77
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	46.47 (4.0)	0.94	3.00	5.00
i/t/a-Merr	45.53 (4.0)		2.06	4.06
PWP	43.47 (3.7)			2.00
SF	41.47 (3.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 44.23 (3.8).

APPENDIX E (CONTINUED)

TABLE 78
JANUARY 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	56.40 (6.0)	4.67	9.47	10.87*
i/t/a-Merr	51.73 (5.2)		4.80	6.20
PWP	46.93 (4.2)			1.40
SF	45.53 (4.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 50.43 (4.8).

* Significance exceeds at .05 level.

TABLE 79
JANUARY 1967 - GRADE III - SPELLING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	PWP	SF
i/t/a-Merr	26.00 (4.6)	0.20	4.13	6.73*
Lipp	25.80 (4.6)		3.93	6.53*
PWP	21.07 (3.9)			2.60
SF	19.27 (3.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 23.11 (4.0).

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 80
JANUARY 1967 - GRADE III - WORD READING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	25.80 (3.8)	2.07	3.62	3.93
i/t/a-Merr	23.73 (3.6)		0.53	1.86
PWP	22.20 (3.3)			0.33
SF	21.87 (3.3)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 23.40 (3.5).

TABLE 81
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	i/t/a-Merr	SF
PWP	39.33 (3.3)	0.13	0.26	5.00
Lipp	39.20 (3.3)		0.13	4.87
i/t/a-Merr	39.07 (3.3)			4.74
SF	34.33 (3.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 37.98 (3.2).

TABLE 82
JANUARY 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	46.67 (4.2)	2.47	4.94	11.87*
i/t/a-Merr	44.20 (3.7)		2.47	9.40
PWP	41.73 (3.5)			6.93
SF	34.80 (2.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all low IQ pupils was 41.26 (3.4).

* Significance exceeds at .05 level.

TABLE 83
JANUARY 1967 - GRADE III - SPELLING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	PWP	SF
i/t/a-Merr	22.00 (3.9)	1.47	3.53	5.07
Lipp	20.53 (3.8)		2.06	3.60
PWP	18.47 (3.5)			1.54
SF	16.93 (3.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 19.66 (3.7).

APPENDIX E (CONTINUED)

TABLE 84
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
CELL DATA FOR UNWEIGHTED MEANS ANALYSIS

			IQ LEVELS		
			b_1 (High)	b_2 (Average)	b_3 (Low)
TREATMENTS	a_1 (SF)	n_{ij}	15	15	15
		SX	666	622	515
		SX^2	30144	27198	18457
		SS_{ij}	573.60000	1405.73334	775.33334
	a_2 (Lipp)	n_{ij}	15	17	35
		SX	718	797	1390
		SX^2	36362	38203	59162
		SS_{ij}	1993.73334	837.76471	3959.14286
	a_3 (PWP)	n_{ij}	15	20	37
		SX	668	843	1311
		SX^2	31274	37031	49759
		SS_{ij}	1525.73334	1498.55	3307.08109
	a_4 (i/t/a-Merr)	n_{ij}	18	15	42
		SX	834	683	1528
		SX^2	39319	31741	61190
		SS_{ij}	677.0	641.73334	5599.90477

APPENDIX E (CONTINUED)

TABLE 85
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
CELL MEANS

		IQ LEVELS			
		b_1 (High)	b_2 (Average)	b_3 (Low)	Total
TREATMENTS	a_1 (SF)	44.40	41.47	34.33	120.20
	a_2 (Lipp)	47.87	46.88	39.71	134.46
	a_3 (PWP)	44.53	42.15	35.43	122.11
	a_4 (i/t/a-Merr)	46.33	45.53	36.38	128.24
		183.13	176.03	145.85	505.01

TABLE 86
SUMMARY OF ANALYSIS OF VARIANCE

Source of Variation	SS	df	MS	F
A (Method)	777.39554	3	259.13184	2.60783
B (IQ Level)	3652.27141	2	1826.13570	19.78720**
AB	57.05330	6	9.50888	0.10303
Within cell	22795.31013	247	92.28870	

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 87
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
ALL 259 PUPILS

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	43.36 (3.7)	2.76	2.29	4.17
i/t/a-Merr	40.60 (3.5)		0.53	1.41
SF	40.07 (3.4)			0.88
PWP	39.19 (3.3)			

HIGH IQ

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	47.87 (4.3)	1.54	3.34	3.47
i/t/a-Merr	46.33 (4.0)		1.80	1.93
PWP	44.53 (3.9)			0.13
SF	44.40 (3.8)			

AVERAGE IQ

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	46.88 (4.1)	1.35	4.73	5.41
i/t/a-Merr	45.53 (4.0)		3.38	4.06
PWP	42.15 (3.6)			0.68
SF	41.47 (3.5)			

LOW IQ

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	39.71 (3.4)	3.33	4.28	5.38
i/t/a-Merr	36.38 (3.1)		2.05	2.05
PWP	35.43 (3.1)			1.10
SF	34.33 (3.0)			

---

[1] Raw score means followed by grade equivalent means.

TABLE 88
JANUARY 1967 - GRADE III - WORD MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	239.390	79.796	3.348*
IQ Levels	2	868.540	434.270	18.222**
Interaction	6	80.310	13.380	0.553
Error	168	4066.530	24.205	
Total	179	5254.770		

* Significance exceeds at .05 level.
** Significance exceeds at .01 level.

TABLE 89
JANUARY 1967 - GRADE III - PARAGRAPH MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	484.650	161.550	1.985
IQ Levels	2	1984.870	992.435	12.192**
Interaction	6	121.950	20.325	0.234
Error	168	14041.310	83.579	
Total	179	16632.780		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 90
JANUARY 1967 - GRADE III - WORD STUDY SKILLS
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	2717.910	905.970	6.905**
IQ Levels	2	3233.100	1616.550	12.321**
Interaction	6	213.540	35.590	0.264
Error	168	22616.250	134.621	
Total	169	28780.800		

** Significance exceeds at .01 level.

TABLE 91
JANUARY 1967 - GRADE III - SPELLING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	912.553	304.184	8.334**
IQ Levels	2	583.213	291.607	7.989**
Interaction	6	40.701	6.784	0.181
Error	168	6310.398	37.562	
Total	169	7846.865		

** Significance exceeds at .01 level.

TABLE 92
GRADE III - IQ
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	72.900	24.300	0.562
IQ Levels	2	24420.000	12210.000	282.472**
Interaction	6	217.200	36.200	0.838
Error	168	7261.900	43.226	
Total	179	31972.000		

** Significance exceeds at .01 level.

TABLE 93
GRADE III - TEACHER EFFECTIVENESS
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	40.151	13.384	1.259
IQ Levels	2	31.111	15.556	1.464
Interaction	6	41.736	6.956	0.6466
Error	168	1807.331	10.758	
Total	169	1920.329		

APPENDIX E (CONTINUED)

TABLE 94
MAY 1967 - WORD MEANING
COMPARISON OF MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	SF	PWP
i/t/a-Merr	28.16 (4.2)	.47	.83	.83
Lipp	27.69 (4.2)		.36	.36
SF	27.33 (4.0)			.00
PWP	27.33 (4.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 27.63 (4.2) with a standard deviation of 5.04.

TABLE 95
MAY 1967 - GRADE III - WORD MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	20.510	6.837	0.349
IQ Levels	2	1117.810	558.905	28.074**
Interaction	6	63.210	10.535	0.529
Error	168	3344.530	19.908	
Total	179	4546.060		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 96
MAY 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	47.31 (4.1)	2.02	2.31	2.35
Lipp	45.29 (3.9)		.29	.33
PWP	45.00 (3.9)			.04
SF	44.96 (3.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 45.64 (4.0) with a standard deviation of 9.56.

TABLE 97
MAY 1967 - GRADE III - PARAGRAPH MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	170.720	56.907	0.763
IQ Levels	2	2932.010	1466.005	18.709
Interaction	6	96.680	16.113	0.206
Error	168	13164.120	78.358	
Total	179	16363.530		

APPENDIX E (CONTINUED)

TABLE 98
MAY 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	53.91 (5.6)	.75	5.55	8.75**
Lipp	53.16 (5.4)		4.80	8.00**
PWP	48.36 (4.5)			3.20
SF	45.16 (3.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 50.14 (4.8) with a standard deviation of 11.57.

** Significance exceeds at .01 level.

TABLE 99
MAY 1967 - GRADE III - WORD STUDY SKILLS
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	2310.470	770.157	7.029**
IQ Levels	2	3010.430	1505.215	13.738**
Interaction	6	222.180	37.030	0.338
Error	168	18407.190	109.567	
Total	179	23950.170		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 100
MAY 1967 - GRADE III - SPELLING
COMPARISON OF MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	PWP	SF
i/t/a-Merr	25.80 (4.6)	.29	1.47	2.87
Lipp	25.51 (4.6)		1.18	2.58
PWP	24.33 (4.2)			1.40
SF	22.93 (4.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 24.64 (4.4) with a standard deviation of 5.61.

TABLE 101
MAY 1967 - GRADE III - SPELLING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	229.990	76.663	2.547
IQ Levels	2	330.000	165.000	5.483**
Interaction	6	25.910	4.318	0.144
Error	168	5055.330	30.091	
Total	179	5641.230		

** Significance exceeds at .01 level.

TABLE 102
MAY 1967 - GRADE III - LANGUAGE
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	54.42 (4.6)	2.06	4.09	4.40
i/t/a-Merr	52.36 (4.4)		2.03	2.34
PWP	50.33 (4.1)			.31
SF	50.02 (4.1)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 51.78 (4.4) with a standard deviation of 10.00.

TABLE 103
MAY 1967 - GRADE III - LANGUAGE
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	562.290	187.430	2.302
IQ Levels	2	3462.940	1731.470	21.271**
Interaction	6	185.750	30.958	0.380
Error	168	13675.590	81.402	
Total	179	17886.570		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 104
JUNE 1967 - GRADE III - WORD MEANING
COMPARISON OF ADJUSTED MEANS

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	22.48 (5.1)	1.94	2.75*	3.76*
i/t/a-Merr	20.55 (4.9)		.82	1.82
SF	19.73 (4.7)			1.00
PWP	18.73 (4.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 20.37 (4.7) with a standard deviation of 6.44.

* Significance exceeds at .05 level.

TABLE 105
JUNE 1967 - GRADE III - WORD MEANING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	340.657	113.552	3.733*
IQ Levels	2	100.362	50.181	1.650
Interaction	6	83.861	13.976	0.460
Error	166	5048.994	30.416	
Total	177	5573.873		

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 106
JUNE 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF ADJUSTED MEANS

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	34.10 (4.9)	3.27	4.42*	5.12*
i/t/a-Merr	30.83 (4.6)		1.15	1.85
PWP	29.68 (4.4)			.70
SF	28.99 (4.3)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 30.90 (4.6) with a standard deviation of 9.63.

* Significance exceeds at .05 level.

TABLE 107
JUNE 1967 - GRADE III - PARAGRAPH MEANING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	691.691	230.564	3.378*
IQ Levels	2	186.312	93.156	1.365
Interaction	6	232.791	38.799	0.569
Error	166	11329.861	68.252	
Total	177	12440.655		

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 108
JUNE 1967 - GRADE III - SPELLING
COMPARISON OF ADJUSTED MEANS

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	33.27 (4.8)	4.04	4.98*	5.29*
i/t/a-Merr	29.23 (4.5)		.93	1.25
SF	28.29 (4.4)			.31
PWP	27.98 (4.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 29.69 (4.6) with a standard deviation of 10.62.

* Significance exceeds at .05 level.

TABLE 109
JUNE 1967 - GRADE III - SPELLING
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	803.071	267.690	2.766*
IQ Levels	2	489.365	244.683	2.529
Interaction	6	143.927	23.988	0.248
Error	166	16063.624	96.769	
Total	177	17499.987		

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 110
JUNE 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF ADJUSTED MEANS

	Means[1]	Differences		
		Lipp	PWP	SF
i/t/a-Merr	46.37 (5.8)	.63	3.77	9.28**
Lipp	45.73 (5.8)		3.13	8.64**
PWP	42.60 (5.3)			5.51**
SF	37.09 (4.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 42.94 (5.3) with a standard deviation of 12.15.

* Significance exceeds at .01 level.

TABLE 111
JUNE 1967 - GRADE III - WORD STUDY SKILLS
ANALYSIS OF COVARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	2398.422	799.474	7.158**
IQ Levels	2	195.372	97.686	0.874
Interaction	6	398.851	66.475	0.595
Error	166	18539.809	111.686	
Total	177	21532.454		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 112
APRIL 1967 - GRADE III - GILMORE ACCURACY
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	37.83	2.05	4.05	6.89
Lipp	35.78		2.00	4.84
PWP	33.78			2.84
SF	30.94			

[1] The grand mean for all 72 pupils in the subsample was 34.58 with a standard deviation of 8.52.

TABLE 113
APRIL 1967 - GRADE III - GILMORE ACCURACY
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	465.828	155.276	2.609
IQ Levels	2	592.579	296.290	4.977**
Interaction	6	527.424	87.904	1.477
Error	60	3571.666	59.528	
Total	71	5157.497		

** Significance exceeds at .01 level.

TABLE 114
APRIL 1967 - GRADE III - GILMORE COMPREHENSION
COMPARISON OF MEANS (UNADJUSTED)

	Means[1]	Differences		
		PWP	Lipp	SF
i/t/a-Merr	26.67	3.06	3.56*	3.73*
PWP	23.61		.50	1.67
Lipp	23.11			.17
SF	22.94			

[1] The grand mean for all 72 pupils in the subsample was 24.08 with a standard deviation of 4.93.

* Significance exceeds at .05 level.

TABLE 115
APRIL 1967 - GRADE III - GILMORE COMPREHENSION
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	164.498	54.833	3.279*
IQ Levels	2	292.748	146.374	8.753**
Interaction	6	262.921	43.820	2.621
Error	60	1003.331	16.722	
Total	71	1723.498		

* Significance exceeds at .05 level.

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 116
APRIL 1967 - GRADE III - GILMORE RATE
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	i/t/a-Merr	PWP
SF	124.44	2.66	4.61	6.16
Lipp	121.78		1.95	3.50
i/t/a-Merr	119.83			1.55
PWP	118.28			

---

[1] The grand mean for all 72 pupils in the subsample was 121.08 with a standard deviation of 23.57.

TABLE 117
APRIL 1967 - GRADE III - GILMORE RATE
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	381.800	127.267	0.214
IQ Levels	2	1058.300	529.150	0.891
Interaction	6	2350.400	391.733	0.659
Error	60	35653.000	594.216	
Total	71	39443.500		

TABLE 118
APRIL 1967 - GRADE III - GATES WORD LIST
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	34.89	.95	2.11	3.17
Lipp	33.94		1.16	2.22
PWP	32.78			1.06
SF	31.72			

[1] The grand mean for all 72 pupils in the subsample was 33.33 with a standard deviation of 4.87.

TABLE 119
APRIL 1967 - GRADE III - GATES WORD LIST
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	102.552	34.184	1.482
IQ Levels	2	104.247	52.124	2.298
Interaction	6	95.869	15.978	0.693
Error	60	1383.331	23.056	
Total	71	1685.999		

APPENDIX E (CONTINUED)

TABLE 120
JUNE 1967 - GRADE III - WORD MEANING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	i/t/a-Merr	PWP
Lipp	25.93 (5.8)	3.33	3.73	5.33
SF	22.60 (5.2)		.40	2.00
i/t/a-Merr	22.20 (5.1)			1.60
PWP	20.60 (4.9)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 22.83 (5.2).

TABLE 121
JUNE 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	40.93 (6.0)	7.26	7.93	8.93*
i/t/a-Merr	33.67 (4.9)		.67	1.67
SF	33.00 (4.8)			1.00
PWP	32.00 (4.7)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 34.90 (5.0).

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 122
JUNE 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	i/t/a-Merr	SF
Lipp	47.53 (6.2)	.06	.13	3.53
PWP	47.47 (6.0)		.07	3.47
i/t/a-Merr	47.40 (6.0)			3.40
SF	44.00 (5.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 46.60 (6.0).

TABLE 123
JUNE 1967 - GRADE III - SPELLING
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	i/t/a-Merr	PWP
Lipp	37.73 (5.4)	4.13	6.60	7.46
SF	33.60 (4.9)		2.47	3.33
i/t/a-Merr	31.13 (4.6)			.86
PWP	30.27 (4.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 high IQ pupils was 33.18 (4.8).

TABLE 124
APRIL 1967 - GRADE III - GILMORE ACCURACY
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	43.67	.34	9.17	11.17*
i/t/a-Merr	43.33		8.83	10.83*
PWP	34.50			2.00
SF	32.50			

[1] The grand mean for all 24 high IQ pupils in the subsample was 38.50.

* Significance exceeds at .05 level.

TABLE 125
APRIL 1967 - GRADE III - GILMORE COMPREHENSION
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	30.00	.33	6.17*	6.50*
i/t/a-Merr	29.67		5.84*	6.17*
PWP	23.83			.33
SF	23.50			

[1] The grand mean for all 24 high IQ pupils in the subsample was 26.75.

*Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 126
APRIL 1967 - GRADE III - GILMORE RATE
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		PWP	Lipp	i/t/a-Merr
SF	132.50	8.83	12.33	12.83
PWP	123.67		3.50	4.00
Lipp	120.17			.50
i/t/a-Merr	119.67			

[1] The grand mean for all 24 high IQ pupils in the subsample was 124.00.

TABLE 127
APRIL 1967 - GRADE III - GATES WORD LIST
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	PWP	SF
i/t/a-Merr	38.00	3.17	4.17	4.83
Lipp	34.83		1.00	1.66
PWP	33.83			.66
SF	33.17			

[1] The grand mean for all 24 high IQ pupils in the subsample was 34.96.

TABLE 128
JUNE 1967 - GRADE III - WORD MEANING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	i/t/a-Merr	PWP
Lipp	22.73 (5.2)	.86	.93	3.73
SF	21.87 (5.1)		.07	2.87
i/t/a-Merr	21.80 (5.1)			2.80
PWP	19.00 (4.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 21.35 (4.9).

TABLE 129
JUNE 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	33.93 (4.9)	.60	2.06	2.73
i/t/a-Merr	33.33 (4.8)		1.46	2.13
SF	31.87 (4.7)			.67
PWP	31.20 (4.6)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 32.58 (4.8).

APPENDIX E (CONTINUED)

TABLE 130
JUNE 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	50.33 (6.5)	1.93	7.80	12.33*
i/t/a-Merr	48.40 (6.2)		5.87	10.40
PWP	42.53 (5.3)			4.53
SF	38.00 (4.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 44.82 (5.7).

* Significance exceeds at .05 level.

TABLE 131
JUNE 1967 - GRADE III - SPELLING
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

	Means[1]	Differences		
		i/t/a-Merr	PWP	SF
Lipp	35.27 (5.1)	3.14	5.34	6.40
i/t/a-Merr	32.13 (4.7)		2.20	3.26
PWP	29.93 (4.6)			1.06
SF	28.87 (4.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 average IQ pupils was 31.55 (3.6).

APPENDIX E (CONTINUED)

TABLE 132
APRIL 1967 - GRADE III - GILMORE ACCURACY
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	Lipp	SF
i/t/a-Merr	38.17	2.00	6.34	10.17
PWP	36.17		4.34	8.17
Lipp	31.83			3.83
SF	28.00			

[1] The grand mean for all 24 average IQ pupils in the subsample was 33.54.

TABLE 133
APRIL 1967 - GRADE III - GILMORE COMPREHENSION
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	SF	Lipp
i/t/a-Merr	26.33	2.50	3.00	5.33
PWP	23.83		.50	2.83
SF	23.33			2.33
Lipp	21.00			

[1] The grand means for all 24 average IQ pupils in the subsample was 23.63.

TABLE 134
APRIL 1967 - GRADE III - GILMORE RATE
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	SF	Lipp
i/t/a-Merr	129.67	4.17	8.84	11.34
PWP	125.50		4.67	7.17
SF	120.83			2.50
Lipp	118.33			

[1] The grand mean for all 24 average IQ pupils in the subsample was 123.58.

TABLE 135
APRIL 1967 - GRADE III - GATES WORD LIST
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	Lipp	SF
i/t/a-Merr	35.00	1.17	2.33	4.67
PWP	33.83		1.16	3.50
Lipp	32.67			2.34
SF	30.33			

[1] The grand mean for all 24 average IQ pupils in the subsample was 32.96.

APPENDIX E (CONTINUED)

TABLE 136
JUNE 1967 - GRADE III - WORD MEANING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	19.40 (4.6)	2.20	3.67	4.00
i/t/a-Merr	17.20 (4.1)		1.47	1.80
PWP	15.73 (3.9)			.33
SF	15.40 (3.8)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 16.93 (4.1).

TABLE 137
JUNE 1967 - GRADE III - PARAGRAPH MEANING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	28.20 (4.2)	3.27	3.40	5.27
i/t/a-Merr	24.93 (3.9)		.13	2.00
PWP	24.80 (3.9)			1.87
SF	22.93 (3.7)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 25.22 (3.9).

APPENDIX E (CONTINUED)

TABLE 138
JUNE 1967 - GRADE III - WORD STUDY SKILLS
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	40.87 (5.0)	.47	3.60	9.74
Lipp	40.40 (4.8)		3.13	9.27
PWP	37.27 (4.4)			6.14
SF	31.13 (3.4)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 37.42 (4.4).

TABLE 139
JUNE 1967 - GRADE III - SPELLING
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	i/t/a-Merr	SF
Lipp	27.27 (4.3)	3.74	3.87	4.07
PWP	23.53 (4.0)		.13	.33
i/t/a-Merr	23.40 (4.0)			.20
SF	23.20 (4.0)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 60 low IQ pupils was 24.35 (4.0).

APPENDIX E (CONTINUED)

TABLE 140
APRIL 1967 - GRADE III - GILMORE ACCURACY
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences i/t/a-Merr	 Lipp	 PWP
SF	32.33	.33	.50	1.66
i/t/a-Merr	32.00		.17	1.33
Lipp	31.83			1.16
PWP	30.67			

[1] The grand mean for all 24 low IQ pupils in the subsample was 31.71.

TABLE 141
APRIL 1967 - GRADE III - GILMORE COMPREHENSION
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences PWP	 SF	 Lipp
i/t/a-Merr	24.00	.83	2.00	5.67*
PWP	23.17		1.17	4.84
SF	22.00			3.67
Lipp	18.33			

[1] The grand mean for all 24 low IQ pupils in the subsample was 21.88.

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 142
APRIL 1967 - GRADE III - GILMORE RATE
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	i/t/a-Merr	PWP
Lipp	126.83	6.83	16.66	21.16
SF	120.00		9.83	14.33
i/t/a-Merr	110.17			4.50
PWP	105.67			

[1] The grand mean for all 24 low IQ pupils in the subsample was 115.67.

TABLE 143
APRIL 1967 - GRADE III - GATES WORD LIST
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	SF	PWP
Lipp	34.33	2.66	2.66	3.66
i/t/a-Merr	31.67		0.00	1.00
SF	31.67			1.00
PWP	30.67			

[1] The grand mean for all 24 low IQ pupils in the subsample was 32.08.

APPENDIX E (CONTINUED)

TABLE 144
JUNE 1967 - GRADE III - WORD MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	416.595	138.865	4.015**
IQ Levels	2	1130.346	565.173	16.342**
Interaction	6	76.990	12.832	0.371
Error	168	5810.133	34.584	
Total	179	7434.064		

** Significance exceeds at .01 level.

TABLE 145
JUNE 1967 - GRADE III - PARAGRAPH MEANING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	770.770	256.923	3.455*
IQ Levels	2	3068.030	1534.015	20.628**
Interaction	6	266.060	44.343	0.596
Error	168	12493.330	74.365	
Total	179	16598.190		

* Significance exceeds at .05 level.
** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 146
JUNE 1967 - GRADE III - WORD STUDY SKILLS
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatment	3	1996.470	665.490	5.301**
IQ Levels	2	2845.480	1422.740	11.332**
Interaction	6	467.780	77.963	0.621
Error	168	21091.720	125.546	
Total	179	26401.450		

** Significance exceeds at .01 level.

TABLE 147
JUNE 1967 - GRADE III - SPELLING
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatment	3	856.010	285.337	2.904*
IQ Levels	2	2650.700	1325.350	13.487**
Interaction	6	178.280	29.713	0.302
Error	168	16509.190	98.269	
Total	179	20194.180		

* Significance exceeds at .05 level.
** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 148
APRIL 1967 - GRADE III - GILMORE ACCURACY
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatment	3	465.828	155.276	2.609
IQ Levels	2	592.579	296.288	4.977**
Interaction	6	527.424	87.904	1.477
Error	60	3571.666	59.528	
Total	71	5157.497		

** Significance exceeds at .01 level.

TABLE 149
APRIL 1967 - GRADE III - GILMORE COMPREHENSION
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatment	3	164.498	54.833	3.279*
IQ Levels	2	292.748	146.374	8.753**
Interaction	6	262.921	43.820	2.621
Error	60	1003.331	16.722	
Total	71	1723.498		

* Significance exceeds at .05 level.
** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 150
APRIL 1967 - GRADE III - GILMORE RATE
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	381.800	127.267	0.214
IQ Levels	2	1058.300	529.150	0.891
Interaction	6	2350.400	391.733	0.659
Error	60	35653.000	594.217	
Total	71	39443.500		

TABLE 151
APRIL 1967 - GRADE III - GATES WORD LIST
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	102.552	34.184	1.483
IQ Levels	2	104.247	52.124	2.298
Interaction	6	95.869	15.978	0.693
Error	60	1383.331	23.056	
Total	71	1685.999		

TABLE 152
MAY 1967 - GRADE III - SAN DIEGO ATTITUDE SCALE COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	Lipp	PWP	SF
i/t/a-Merr	19.18	.09	.29	2.09
Lipp	19.09		.20	2.00
PWP	18.89			1.80
SF	17.09			

[1] The grand mean for all 180 pupils was 18.56 with a standard deviation of 4.72.

TABLE 153
MAY 1967 - GRADE III - SAN DIEGO ATTITUDE SCALE COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	PWP	SF
Lipp	20.33	.06	1.40	2.66
i/t/a-Merr	20.27		1.34	2.60
PWP	18.93			1.26
SF	17.67			

[1] The grand mean for all 60 high IQ pupils was 19.30.

APPENDIX E (CONTINUED)

TABLE 154
MAY 1967 - GRADE III - SAN DIEGO ATTITUDE SCALE
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	i/t/a-Merr	Lipp	SF
PWP	20.40	.13	2.07	3.53
i/t/a-Merr	20.27		1.94	3.40
Lipp	18.33			1.46
SF	16.87			

[1] The grand mean for all 60 average IQ pupils was 18.97.

TABLE 155
MAY 1967 - GRADE III - SAN DIEGO ATTITUDE SCALE
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	i/t/a-Merr	SF
Lipp	18.60	1.27	1.60	1.87
PWP	17.33		.33	.60
i/t/a-Merr	17.00			.27
SF	16.73			

[1] The grand mean for all 60 low IQ pupils was 17.42.

APPENDIX E (CONTINUED)

TABLE 156
MAY 1967 - GRADE III - SAN DIEGO ATTITUDE SCALE
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	132.018	44.006	2.09
IQ Levels	2	121.213	60.607	2.803*
Interaction	6	99.102	16.517	0.764
Error	168	3631.998	21.619	
Total	179	3984.321		

* Significance exceeds at .05 level.

TABLE 157
FEBRUARY 1967 - GRADE III - BOOKS READ
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	PWP	i/t/a-Merr
Lipp	10.77	.46	4.56**	5.59**
SF	10.31		4.10*	5.13**
PWP	6.21			1.03
i/t/a-Merr	5.18			

[1] The grand mean for all 180 pupils was 8.12 with a standard deviation of 6.76.

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

APPENDIX E (CONTINUED)

TABLE 158
FEBRUARY 1967 - GRADE III - BOOKS READ
COMPARISON OF HIGH IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	PWP	i/t/a-Merr
Lipp	12.36	0.65	5.69	7.43**
SF	11.71		5.04	6.78*
PWP	6.67			1.74
i/t/a-Merr	4.93			

[1] The grand mean for all 60 high IQ pupils was 8.89.

** Significance exceeds at .01 level.

* Significance exceeds at .05 level.

TABLE 159
FEBRUARY 1967 - GRADE III -- BOOKS READ
COMPARISON OF AVERAGE IQ MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	PWP	i/t/a-Merr
Lipp	10.00	0.00	3.87	4.73
SF	10.00		3.87	4.73
PWP	6.13			0.86
i/t/a-Merr	5.27			

[1] The grand mean for all 60 average IQ pupils was 7.85.

APPENDIX E (CONTINUED)

TABLE 160
FEBRUARY 1967 - GRADE III - BOOKS READ
COMPARISON OF LOW IQ MEANS (UNADJUSTED)

	Means[1]	Differences SF	 PWP	 i/t/a-Merr
Lipp	10.00	0.79	4.15	4.67
SF	9.21		3.36	3.88
PWP	5.85			0.52
i/t/a-Merr	5.33			

[1] The grand mean for all 60 low IQ pupils was 7.60.

TABLE 161
FEBRUARY 1967 - GRADE III - BOOKS READ
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	1055.849	351.950	8.532**
IQ Levels	2	54.202	27.101	0.657
Interaction	6	50.890	8.482	0.206
Error	168	6682.387	41.249	
Total	179	7843.328		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 162

MAY 1967 - GRADE III - SILENT READING ACHIEVEMENT AND READING EXPECTANCY BY CLASSES[1]

Program	Teacher	No. of Pupils at Least 1/2 Grade Above[2]	Total Reading Study Students	50% or More Achieved at Least 1/2 Grade Above Predicted Score
SF	A	10	15	X
	B	0	3	
	C	5	16	
	D	7	16	
Lipp	E	4	13	X[3]
	F	7	12	X[3]
	G	13	18	
	H	2	11	X[3]
	I	10	15	
SF + PWP	J	4	12	
	K	3	10	
	L	3	23	X
	M	6	11	X[3]
	N	10	14	X[3]
i/t/a-Merr	O	11	15	
	P	5	13	
	Q	3	16	X[3]
	R	16	20	
	S	4	11	
Totals		123	264	8

---

[1] Comparisons made between predicted achievement on Bond and Tinker Reading Grade Scores and actual achievement on Stanford Achievement Test, Primary II, Form X.

[2] Number includes those who were one-half grade level above the Bond and Tinker Reading Grade Score on at least two of three reading subtests (Word Meaning, Paragraph Meaning, Word Study Skills - Stanford Achievement Test, Primary II, Form X).

[3] Indicates classes which also qualified as over-achieving in Grades I and II.

APPENDIX E (CONTINUED)

TABLE 163

TEACHING EFFECTIVENESS[1] AND PUPIL ACHIEVEMENT IN GRADE III

Program	Teacher[2]	Plan	Know	Com	Manage	Pers	Obj	Totals[3]
SF	A	2.92	3.08	2.83	3.00	3.45	2.92	18.20
Lipp	F	2.42	2.25	2.33	2.25	2.55	2.42	14.22
Lipp	G	2.42	2.67	2.92	2.83	2.67	2.75	16.26
Lipp	I	3.67	3.50	3.42	3.08	3.67	3.50	20.84
SF + PWP	M	2.00	2.08	2.33	2.25	2.42	2.17	13.25
SF + PWP	N	2.42	2.25	2.33	2.50	2.75	2.42	14.67
i/t/a-Merr	O	3.25	3.17	3.08	3.17	3.58	3.17	19.42
i/t/a-Merr	R	2.83	3.17	2.92	3.00	3.00	2.92	17.84
Mean[4]		2.74	2.77	2.77	2.76	3.01	2.78	16.84
Mean[5]		2.02	2.12	2.10	2.15	2.32	2.08	12.79
Total Mean[6]		2.33	2.39	2.38	2.41	2.61	2.37	14.49

---

[1] Defined by the Hayes Teacher Rating Scale on a 0-4 scale where 2 is average.

[2] Identified in Table 162 and for whom the majority of pupils achieved one-half level above expectancy levels.

[3] Based upon 0-24 point scale.

[4] Scores attained by teachers who had 50% or more reading study students at least one-half grade level above expectancy levels.

[5] Scores attained by teachers who did not have 50% or more reading study students at least one-half grade level above expectancy levels.

[6] Scores attained by all nineteen reading study teachers in Grade III.

APPENDIX E (CONTINUED)
TABLE 164
JUNE 1967 - GRADE III - SILENT READING ACHIEVEMENT AND READING EXPECTANCY BY CLASSES[1]

Program	Teacher	No. of Pupils at Least 1/2 Grade Above[2]	Total Reading Study Students	50% or More Achieved at Least 1/2 Grade Above Predicted Score
SF	A	7	15	
	B	1	3	
	C	9	16	X
	D	10	16	X
Lipp	E	8	13	X
	F	9	12	X[3]
	G	17	18	X[3]
	H	5	11	
	I	11	15	X[3]
SF + PWP	J	6	12	X
	K	4	10	
	L	5	23	
	M	4	11	
	N	12	14	X[3]
i/t/a-Merr	O	13	15	X[3]
	P	5	13	
	Q	9	16	X
	R	16	20	X[3]
	S	5	11	
Totals		156	264	11

[1] Comparisons made between predicted achievement on Bond and Tinker Reading Grade Scores and actual achievement on Stanford Achievement Test, Intermediate I, Form W.

[2] Number includes those who were one-half grade level above the Bond and Tinker Reading Grade Score on at least two of three reading subtests (Word Meaning, Paragraph Meaning, Word Study Skills - Stanford Achievement Test, Intermediate I, Form W).

[3] Indicates classes which also qualified on Stanford Achievement Test, Primary II, Form X (See Table 162).

APPENDIX E (CONTINUED)

TABLE 165

TEACHING EFFECTIVENESS[1] AND PUPIL ACHIEVEMENT IN GRADE III

Program	Teacher[2]	Plan	Know	Com	Manage	Pers	Obj	Totals[3]
SF	C	2.67	2.45	2.58	2.67	3.42	2.67	16.46
SF	D	1.92	2.00	1.64	2.08	2.00	2.00	11.64
Lipp	E	1.33	1.33	1.67	1.75	1.73	1.42	9.23
Lipp	F	2.42	2.25	2.33	2.25	2.55	2.42	14.22
Lipp	G	2.42	2.67	2.92	2.83	2.67	2.75	16.26
Lipp	I	3.67	3.50	3.42	3.08	3.67	3.50	20.84
SF + PWP	J	2.58	2.58	2.75	2.83	3.17	2.67	16.58
SF + PWP	N	2.42	2.25	2.33	2.50	2.75	2.42	14.67
i/t/a-Merr	O	3.25	3.17	3.08	3.17	3.58	3.17	19.42
i/t/a-Merr	Q	1.75	2.42	1.92	1.75	1.92	1.83	11.59
i/t/a-Merr	R	2.83	3.17	2.92	3.00	3.00	2.92	17.84
Mean[4]		2.48	2.53	2.51	2.54	2.77	2.52	15.34
Mean[5]		2.12	2.21	2.22	2.23	2.40	2.17	13.35
Total Mean[6]		2.33	2.39	2.38	2.41	2.61	2.37	14.49

[1] Defined by the Hayes Teacher Rating Scale on a 0-4 scale where 2 is average.

[2] Identified in Table 164 and for whom the majority of pupils achieved one-half level above expectancy levels.

[3] Based upon 0-24 point scale.

[4] Scores attained by teachers who had 50% or more reading study students at least one-half grade level above expectancy levels.

[5] Scores attained by teachers who did not have 50% or more reading study students at least one-half grade level above expectancy levels.

[6] Scores attained by all nineteen reading study teachers in Grade III.

APPENDIX E (CONTINUED)

TABLE 166

PUPIL ATTITUDE[1] IN GRADE III

Program	Teacher[2]	N. Positive Attitude[3]	Total Reading Study Students	50% or More Positive Attitude[4]
SF	A	4	15	
	B	3	3	X[5]
	C	5	16	
	D	9	16	X
Lipp	E	7	13	X
	F	11	12	X[5]
	G	12	18	X[5]
	H	3	11	
	I	8	15	X[5]
SF + PWP	J	8	12	X[5]
	K	6	10	X
	L	13	23	X[6]
	M	9	11	X[5]
	N	13	14	X[5]
i/t/a-Merr	O	10	15	X[5]
	P	6	13	
	Q	10	16	X
	R	16	20	X[5]
	S	4	11	
Totals		157	264	14

[1] Measured by San Diego County Inventory of Reading Attitude.

[2] Letter designations are for the same teachers as in Tables 164 and 165.

[3] Number of reading study students in each classroom with stanine score of six or better (raw score 19+) on San Diego County Inventory of Reading Attitude.

[4] Classes where 50% or more of students had above average attitudes toward reading.

[5] These classes also qualified as having above average attitudes in Grades I and II.

[6] This class also qualified as having an above average attitude in Grade II.

APPENDIX E (CONTINUED)

TABLE 167

TEACHING EFFECTIVENESS[1] AND PUPIL ATTITUDE[2] IN GRADE III

Program	Teacher[3]	Plan	Know	Com	Manage	Pers	Obj	Total
SF	B	2.50	2.58	2.58	2.50	2.75	2.50	15.41
	D	1.92	2.00	1.64	2.08	2.00	2.00	11.64
Lipp	E	1.33	1.33	1.67	1.75	1.73	1.42	9.23
	F	2.42	2.25	2.33	2.25	2.55	2.42	14.22
	G	2.42	2.67	2.92	2.83	2.67	2.75	16.26
	I	3.67	3.50	3.42	3.08	3.67	3.50	20.84
SF + PWP	J	2.58	2.58	2.75	2.83	3.17	2.67	16.58
	K	2.25	2.50	2.58	2.58	2.75	2.42	15.08
	L	2.50	2.30	2.58	2.33	2.75	2.33	14.79
	M	2.00	2.08	2.33	2.25	2.42	2.17	13.25
	N	2.42	2.25	2.33	2.50	2.75	2.42	14.67
i/t/a-Merr	O	3.25	3.17	3.08	3.17	3.58	3.17	19.42
	Q	1.75	2.42	1.92	1.75	1.92	1.83	11.59
	R	2.83	3.17	2.92	3.00	3.00	2.92	17.84
Mean[4]		2.42	2.49	2.50	2.49	2.69	2.47	15.06
Mean[5]		2.13	2.14	2.05	2.17	2.39	2.12	13.00
Total Mean[6]		2.33	2.39	2.38	2.41	2.61	2.37	14.49

[1] Teaching Effectiveness as measured by Hayes Teacher Rating Scale.

[2] Pupil Attitude as measured by San Diego County Inventory of Reading Attitude.

[3] Letter designations are for the same teachers in Tables 164, 165, and 166. These are teachers in whose rooms 50% or more of the reading study students scored above average (stanine six, raw score 19+) on the San Diego County Inventory of Reading Attitudes.

[4] Scores attained by teachers whose classes qualified as having a better than average attitude toward reading.

[5] Scores attained by teachers whose classes did not qualify as having a better than average attitude toward reading.

[6] Scores attained by all nineteen reading study teachers in Grade III.

TABLE 168
MAY 1967 - GRADE III - SCIENCE AND SOCIAL STUDIES
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	SF	PWP	Lipp
i/t/a-Merr	25.38 (4.3)	.27	.31	.62
SF	25.11 (4.3)		.04	.35
PWP	25.07 (4.3)			.31
Lipp	24.76 (4.3)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 25.08 (4.3) with a standard deviation of 5.08.

TABLE 169
MAY 1967 - GRADE III - SCIENCE AND SOCIAL STUDIES
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	8.780	2.927	0.134
IQ Levels	2	822.340	411.170	18.827**
Interaction	6	126.860	21.143	0.968
Error	168	3668.930	21.839	
Total	179	4626.910		

** Significance exceeds at .01 level.

TABLE 170
MAY 1967 -GRADE III - ARITHMETIC COMPUTATION
COMPARISON OF MEANS (UNADJUSTED)

		Differences		
	Means[1]	PWP	i/t/a-Merr	Lipp
SF	43.16 (4.3)	1.63	1.74	1.83
PWP	41.53 (4.2)		.11	.20
i/t/a-Merr	41.42 (4.1)			.09
Lipp	41.33 (4.1)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 41.86 (4.2) with a standard deviation of 8.48.

TABLE 171
MAY 1967 - GRADE III - ARITHMETIC COMPUTATION
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	101.450	33.817	0.521
IQ Levels	2	1258.710	629.355	9.696**
Interaction	6	613.130	102.188	1.574
Error	168	10904.250	64.906	
Total	179	12877.540		

** Significance exceeds at .01 level.

TABLE 172
MAY 1967 - GRADE III - ARITHMETIC CONCEPTS
COMPARISON OF MEANS (UNADJUSTED)

	Means[1]	Differences		
		Lipp	i/t/a-Merr	PWP
SF	35.09 (4.7)	.07	.38	1.31
Lipp	35.02 (4.7)		.31	1.24
i/t/a-Merr	34.71 (4.7)			.93
PWP	33.78 (4.5)			

[1] Raw score means followed by grade equivalent means. The grand mean for all 180 pupils was 34.65 with a standard deviation of 6.48.

TABLE 173
MAY 1967 - GRADE III - ARITHMETIC CONCEPTS
ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Treatments	3	49.310	16.437	0.474
IQ Levels	2	1566.230	783.115	22.579**
Interaction	6	72.630	12.105	0.349
Error	168	5826.780	34.683	
Total	179	7514.950		

** Significance exceeds at .01 level.

APPENDIX E (CONTINUED)

TABLE 174
1966-67 - GRADE III - CORRELATION MATRIX A

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. January Word Meaning																							
2. January Paragraph Meaning	80																						
3. January Word Study Skills	67	67																					
4. January Spelling	74	73	70																				
5. Books Read	27	13	21	12																			
6. IQ	46	42	37	28	05																		
7. Teacher Effectiveness Rating	24	25	28	20	02	12																	
8. May Word Meaning	76	79	63	67	12	54	21																
9. May Paragraph Meaning	73	81	63	66	17	49	21	84															
10. May Word Study Skills	64	62	89	69	17	37	29	62	61														
11. May Language	73	73	74	66	30	50	30	68	68	71													
12. May Spelling	66	70	63	87	22	19	15	62	66	60	64												
13. May Science and Social Studies	60	50	42	33	11	52	19	61	55	45	50	27											
14. May Arithmetic Computation	48	51	45	40	33	38	15	45	53	46	56	46	33										
15. May Arithmetic Concepts	60	66	65	46	20	50	24	66	72	58	66	47	49	57									
16. San Diego Attitude Scale	34	49	22	42	21	15	08	33	35	27	40	43	05	24	21								
17. June Word Meaning	74	73	61	66	22	49	16	76	69	58	70	60	55	51	59	35							
18. June Paragraph Meaning	74	76	66	65	21	49	14	76	76	63	75	61	56	53	62	36	82						
19. June Word Study Skills	68	69	83	76	24	34	30	64	67	86	75	74	42	53	59	33	61	68					
20. June Spelling	69	70	61	72	26	34	18	64	63	57	69	73	34	44	53	40	66	73	65				
21. Gilmore Accuracy	61	62	56	54	02	31	-05	50	52	50	55	47	25	34	40	42	56	65	49	48			
22. Gilmore Comprehension	51	59	48	43	-07	43	-02	47	51	43	53	32	36	29	38	33	53	57	34	42	74		
23. Gilmore Rate	38	39	26	25	44	08	-09	34	36	25	28	29	16	23	26	35	30	31	29	32	13	00	
24. Gates Word List	77	74	68	75	09	33	05	63	57	59	60	61	24	30	48	56	64	50	58	57	64	46	36

APPENDIX F

PUPIL RETENTIONS

(A Study by Linda Nash)

During the first two years of the New Castle Reading Study, a total of twenty percent of the Lippincott pupils were retained in either first or second grade compared to eight percent of the i/t/a-Merrill pupils, almost eight percent of the Scott, Foresman pupils, and eleven percent of the Phonics and Word Power pupils. The major concern is that, while the Lippincott group had the greatest number of retainees in both years of the study, other results showed that they also had the best overall achievement scores.

A case study approach was used in an attempt to view each retainee individually in relation to his own physical, emotional, intellectual, and educational abilities and limitations. By employing a similar format for each, the individual cases and the four treatment groups were studied for their generalities and specifics. The information obtained was taken from that recorded in each student's cumulative record folder on file in the various New Castle schools.

A total of fifty-seven subjects were studied[1], of which thirty-one were retained at the end of the 1964-65 school year in first grade. In addition, at the end of the 1965-66 school year, twelve were retained in second grade and fourteen were retained in first grade. Of the total fifty-seven subjects, forty percent were Lippincott pupils, twenty-eight percent Phonics and Word Power, twenty-three percent i/t/a-Merrill, and nine percent Scott, Foresman. It is also noted that twenty-eight percent of the Lippincott retainees were from one school located in a low socio-economic area of New Castle.

In addition to the individual case study information gathered, twelve of the teachers involved in the retention problem were interviewed. Each teacher was asked the following questions:

1. What is the rationale of the school district regarding retention? of the principal? of the teachers? of yourself?

---

[1] This figure does not represent the exact population of retainees due to the loss of some who moved out of the New Castle School District.

APPENDIX F (CONTINUED)

2. What factors are used as criteria for determining who shall be retained?

3. If you retain a child, do you prefer to keep him in your class next year or move him into another class?

4. Do you feel the method you used to teach reading had any effect upon the children you retained?

5. How is the retainee's academic and social adjustment in his new classroom?

6. Do you feel there is any value in a kindergarten experience?

According to the teachers interviewed, the New Castle School District allows each school principal certain discretionary powers regarding retention practices. The policies of the various schools are for the most part similar. After the next to the final report card period ends, the principal checks with teachers regarding possible retentions. At this time parents are usually requested to attend a conference with the principal and teacher. Other school personnel who might be called in on the conference are the guidance counselor, school psychologist, and field director. If the parents strongly oppose the decision, the child is promoted after the parents sign a release slip which frees the teacher from any consequences which might ensue. It was also reported that the New Castle schools generally follow a policy which allows a child to be retained once in the primary grades and once in the intermediate grades. Also, one teacher reported it is usually true that a child must repeat first grade once before being placed into Special Education. She added that some schools allow a teacher to "place" rather than "pass" a child into the next grade. This is recorded on his report card and cumulative record and is done for reasons of age, size, and parental objection to retention.

Immaturity was named most often as a criterion for retention. It was commonly thought among the teachers interviewed that one who is immature (physically, emotionally, socially, or academically), one who lacks the foundation which should have been established in first grade, can genuinely benefit from another year in first grade. However, one who is immature mentally, one who simply does not have the native intelligence to achieve near grade level, will not benefit as much from being retained. Most teachers denied that a criterion would be completion of a specific page or chapter in a book, although achievement test scores (if consistent with the child's overall progress) on reading comprehension subtests are considered. A grade level score of 1.0 in paragraph meaning designates a nonreader whose lack of achievement cannot be overlooked. Teachers suggested the use of January achievement test scores as warning signals for possible retentions. One Lippincott teacher said that consideration must certainly be given to the amount of material covered and learned. Physical size and chronological age are also factors. A child who is older and larger than many of his peers may be "placed" into the next grade due to a possibility of poor social adjustment if retained.

Teachers who wish to keep their own retainees say that they best understand the child's abilities and limitations. Also, they cite the security of a familiar setting and the initial success gained by "knowing the ropes" of a particular situation. Those feeling it best for the child to have a change of scenery say that a new classroom and a new teacher give the child a feeling of a fresh start. Also, they cite the possibility that there may have been a personality conflict between the previous teacher and the child which only added to the retainee's problems. Both sides agree that such a decision for the most part, depends upon the individual child and the school situation.

The importance of considering each child as an individual came up again when the teachers were asked to voice their opinion about the relationship of the reading approach used and the retention problem. The most common answer was that the deciding factor was within the individual rather than within the method used to teach reading. It was felt that the individuals they failed would generally have failed regardless of the manner in which they were taught reading. One teacher did suggest that perhaps a whole-classroom approach, such as the Lippincott, does add to the potential retainee's problems. While not in itself a cause, it could be a contributing factor to an already troubled child. It was pointed out that the potential failure is the one child in the class who, above all others, requires special attention and individual help and guidance for which a whole-classroom approach does not account.

Socially speaking, the retainee adjusts very well to his new classmates. Because he was deemed immature for his own age-group peers, he adjusts well to a younger group of classmates. Academically, the initial success experienced is rewarding, although short-lived. As one teacher so aptly put it, the retainee seems to "coast along" on what he has learned last year, until perhaps December or January, at which time the class once again leaves him behind. Hopefully, however, the retainee has established somewhat of a foundation upon which to build in future years. Several teachers suggested that the greatest adjustment is on the part of the parent. Once the parents accept the decision, the child will be only too glad to have another chance.

The need for a public kindergarten in New Castle is keenly felt by the teachers. The child with a preschool opportunity gains valuable social experience and exposure. This extra bit of readiness is most welcomed by the first-grade teachers. However, such an experience does not have lasting effects upon the child. The teachers interviewed generally agreed that it is but an initial advantage and is social rather than academic. It was said that by November, those with a kindergarten experience are not distinguishable from those without kindergarten.

Regarding the total fifty-seven retainees studied, it was found that the average IQ score obtained from the Pintner-Cunningham Primary Test of Intelligence (1964 revision) was 81 which is classified as low average. This intelligence test was administered at the beginning of the subjects' respective first grades. Using June 9, 1967 as a standard calculation date, the average chronological age of the fifty-seven subjects was eight years, six months, and their average corresponding mental age was seven years. (See Table 1). The average reading readiness score as obtained from the Lee-Clark Reading Readiness Test (administered in May prior to Grade I) was 30.7 which is classified as low and is given a grade level equivalent of .4 (1.0 would indicate an average first grade level). A general comparison between the reading expectancy and reading level scores shows that ninety-five percent of the retainees were underachievers, both at the time of their retention and at the present time. The average degree of retardation at the time of their retention, or the difference between the average reading expectancy and reading level scores, was one year, six months.

It was also found that seventy percent of the parents of the retainees had not finished high school and twenty-six percent of the parents were either separated, divorced, or remarried. Fourteen percent of the mothers worked and ten percent of the families were on public assistance. The majority of parents and children were born and reared in New Castle. (See Table 1).

The several case studies chosen for inclusion in this report are representative of the total fifty-seven instances of retention. Fictitious names are reported to avoid possible embarrassment.

Concerning the case study outline itself, there are several terms, abbreviations, and scores which need to be defined before the case studies can take on their intended meanings.

APPENDIX F (CONTINUED)

1. CA--(chronological age)--The age which appears on the case studies represents the subject's age as of June 9, 1967.

2. MA--(mental age)--This age is computed by multiplying the chronological age by the IQ and dividing that number by 100. The IQ used for this purpose was from the Pintner-Cunningham Primary Test of Intelligence administered at the beginning of the respective first grades of each of the subjects.

3. LCRR--The score given for the Lee Clark Reading Readiness Test is a raw score followed by an interpretation of it in terms of high, average or low. Each raw score also has a corresponding grade level equivalent and some suggest delayed entry as being advisable.

4. IQ--The scores given for the Pintner-Cunningham Primary Test of Intelligence and the Binet are transformed IQ scores.

5. SAT--The Stanford Achievement Test score recorded represents the comparable grade level equivalent on the paragraph meaning subtest. For example, a score of 1.5 represents the average of the test scores obtained from a sample of all the children in the fifth month of grade one in the schools of the nation. For the most part, the reading study tests were administered in April or May of the respective years. Also, it can be pointed out here that those children who entered the study in 1965-66 and were repeating first grade this year (1966-67) were not administered the Stanford Achievement Tests and thus do not have reading level scores during their second year in Grade I.

6. R--(repeated)--The letter (R) after a grade means that the child repeated that particular grade during that particular year.

7. RE--(reading expectancy)--This score, which represents the grade level at which a given child might be expected to score, is computed by multiplying IQ/100 times years in school and adding one year to this total. The IQ score used was from the Pintner-Cunningham Test which was administered at the beginning of the respective first grades of each of the subjects.

APPENDIX F (CONTINUED)

8. RL--(reading level)--This score represents the grade level at which the child is actually reading as shown by the score the child received on the paragraph meaning subtest of the SAT.

9. DR--(degree of retardation)--This score represents the difference between the reading expectancy and reading level scores of a subject. It, too, is a grade level equivalent score.

10. RC--(regular classroom)--The New Castle School District uses the 1955 edition of the Scott, Foresman series in its regular classrooms. The Reading Study uses the 60's edition. The manner of presentation of the material by the teachers is also of considerable difference so as to warrant a qualifying statement about the two approaches being non-comparable.

11. BI--The section marked Background Information is taken directly from teacher comments in the subjects' cumulative record folders.

APPENDIX F (CONTINUED)

TABLE 1
INFORMATION FROM CASE STUDIES

	Column 1						Column 2		Column 3		Column 4			Column 5*				
	Education of Parents						Marital Status		Occupation		Born in New Castle			RL> RE				
	High School		Not High School		High School +		Married & Living Together	Separated	Mother Working	On Public Assist.	Mother	Father	Child		IQ	CA	MA	LCRR
	M	F	M	F	M	F												
TOTAL 100% N=57	27%	27%	70%	70%	3%	3%	74%	26%	14%	10%	61%	58%	83%	5%	81	8-6	7-0	30.7
LIPP 40% N=23	46%	27%	50%	64%	4%	9%	78%	22%	17%	17%	57%	57%	96%	5%	83	8-6	7-2	31.5
PWP 28% N=16	13%	15%	77%	77%	0%	0%	75%	25%	13%	6%	75%	56%	75%	0%	82	8-6	7-0	32.3
i/t/a-MERR 23% N=13	17%	0%	83%	100%	0%	0%	70%	30%	13%	6%	38%	44%	50%	6%	77	8-6	6-8	29.5
SF 9% N=5	50%	0%	50%	100%	0%	0%	60%	40%	0%	20%	80%	60%	100%	20%	79	8-4	6-8	25.4

*See pages F-6 and F-7 for explanation of abbreviations.

APPENDIX F (CONTINED)

Lippincott-A

CASE STUDY OUTLINE

1. Name: Robert

2. Birthday: June 6, 1958
New Castle, Pa.

a. CA - 9-0   b. MA - 7-0

3. Family:

	Birthplace	Occupation	Education	Marital Status	No. of Children
Father	Virginia	On Public Assistance	Special Education	Married and Living Together	3
Mother	Ohio	Housewife	HS Grad		

4. Test Results:

Intelligence Tests

Name of Test	Date	Grade	Result
1. Lee Clark Reading Readiness	5/64	Preschool	21 Low
2. Pintner-Cunningham	9/64	1st	78 IQ
3. Binet	3/65	1st	71 IQ
4. Pintner-Cunningham	11/65	1st (R)	87 IQ
5. Binet	3/66	1st (R)	87 IQ

5. Reading Levels:

Year	Grade	Treatment	RE	RL*	Degree of Retardation
1964-65	1st	Lipp	1.8	1.5	.3
1965-66	1st (R)	RC	2.6	1.4	1.2
1966-67	2nd	RC	3.3	1.6	1.7

* Stanford Achievement Test - Paragraph Meaning Subtest - administered at end of respective school years.

6. Diagnostic Considerations:

   a. Background Information:

      1964-65: Robert's hearing test resulted in a referral to have his tonsils checked. His speech was diagnosed as mildly defective. Robert has a very short attention span.

      1965-66: Robert is a discipline problem. He is bored and not achieving as he should be. He is very immature.

   b. Intellectual:

      1964-65: Robert scored low on his reading readiness test. His score suggested that a year's delay of entry may have been advisable. His Pintner-Cunningham score showed a 78 IQ which is classified as being borderline defective. In March 1965, Robert was administered a Binet on which he scored a 71 IQ which is also borderline defective.

      1965-66: At the beginning of his repeated year in first grade, Robert scored an 87 IQ on another form of the Pintner-Cunningham. This is interpreted as being low average. In March 1966, he was given another form of the Binet on which he also scored and 87 IQ (low average).

   c. Educational:

      1964-65: Robert's SAT score for paragraph meaning showed his reading level to be three months behind his reading expectancy score.

      Robert was recommended for special education because of his Binet score but his parents preferred him to repeat a grade instead.

      1965-66: Robert's SAT score showed his reading level to be one year, two months behind his reading expectancy score. This achievement score was a month behind his 1964-65 score on a comparable testing device.

      Robert was not recommended for special education this year because of the gain on his Binet IQ test.

APPENDIX F (CONTINUED)

1966-67: Robert gained only two months on his reading level score and remained one year, seven months behind his reading expectancy.

Lippincott-B

CASE STUDY OUTLINE

1. Name: David

2. Birthday: June 23, 1958
New Castle, Pa.

A. CA - 9-0 b. MA - 6-8

3. Family:

	Birthplace	Occupation	Education	Marital Status	No. of Children
Father	New Castle	Laborer	HS Grad	Married and Living Together	2
Mother	New Castle	Housewife	HS Grad		

4. Test Results:

Intelligence Tests

	Name of Test	Date	Grade	Result
1.	Lee Clark Reading Readiness	5/64	Preschool	38 Low
2.	Pintner-Cunningham	9/64	1st	75 IQ
3.	Binet	3/66	2nd	80 IQ

5. Reading Levels:

Year	Grade	Treatment	RE	RL*	Degree of Retardation
1964-65	1st	Lipp	1.8	1.4	.4
1965-66	2nd	Lipp	2.5	1.8	.7
1966-67	2nd (R)	RC	3.3	2.8	.5

* Stanford Achievement Test - Paragraph Meaning Subtest - administered at end of respective school years.

6. Diagnostic Considerations:

   a. Background Information:

      1964-65: David passed first grade due to his size and the reading study. He has poor visual-motor coordination and has difficulty verbalizing and organizing ideas. His verbal explanations are very confusing.

      1965-66: David will be retained next year so that he can grasp the basics he missed previously. In spite of his low IQ, he is eager to work and will meet more success once he can acquire a foundation.

   b. Intellectual:

      1964-65: David scored a low average on his preschool reading readiness test which corresponded to a .4 grade level equivalent. His Pintner-Cunningham showed a 75 IQ which is classified as borderline defective.

      1965-66: On a Binet which was administered this year, David scored an 80 IQ which is termed low average.

   c. Education:

      1964-65: David's SAT paragraph meaning scores showed him to be reading at a 1.4 level. As compared with his reading expectancy level, it showed a four-month's degree of retardation.

      1965-66: David gained four months on his reading level this year, and remained seven months behind his reading expectancy.

      1966-67: After repeating second grade, David gained a full year on his reading level and remained five months behind his reading expectancy.

i/t/a-Merrill-A

CASE STUDY OUTLINE

1. Name: Carolyn

2. Birthday: November 11, 1958
Pennsylvania

a. CA - 8-7 b. MA - 7-8

3. Family:

	Birthplace	Occupation	Education	Marital Status	No. of Children
Father	Pennsylvania	Sells Housewares	Not HS Grad	Married and Living Together	
Mother	Pennsylvania	Housewife	Not HS Grad		5

4. Test Results:

Intelligence Tests

Name of Test	Date	Grade	Result
1. Lee Clark Reading Readiness	5/64	Preschool	19 Low
2. Pintner-Cunningham	10/65	1st (R)	91 IQ

5. Reading Levels:

Year	Grade	Treatment	RE	RL*	Degree of Retardation
1964-65	1st	i/t/a-Merr	1.9	1.3	.6
1965-66	1st (R)	i/t/a-Merr	2.8	1.7	1.1
1966-67	2nd	i/t/a-Merr	3.7	1.9	1.8

* Stanford Achievement Test - Paragraph Meaning Subtest - administered at end of respective school years.

APPENDIX F (CONTINUED)

6. Diagnostic Considerations:

   a. Background Information:

      1964-65: Carolyn was advised to stay out for a year. She is often tardy, as she gets up by herself. She never finishes an assignment due to her slowness.

      1965-66: Carolyn was administered a Binet because she was regressing instead of progressing. Results showed that she scored close to average on the vocabulary and comprehension subtests but was below age level on eye-hand motor coordination and distinguishing between similarities and differences. Testing also showed her to have a very short attention span.

   b. Intellectual:

      1964-65: Carolyn's low score on her reading readiness test suggested that she be delayed for a year before entering grade one.

      1965-66: Carolyn scored a 91 IQ on her Pintner-Cunningham. This is interpreted as being normal or average.

   c. Educational:

      1964-65: Carolyn's reading level score showed her to be six months behind her reading expectancy.

      1965-66: Carolyn gained four months on her reading level score and remained one year, one month behind her reading expectancy.

      1966-67: Carolyn gained only two months on her reading level score and remained almost two years behind her reading expectancy.

APPENDIX F (CONTINUED)

i/t/a-Merrill-B

CASE STUDY OUTLINE

1. Name: Larry

2. Birthday: November 30, 1958
New Castle, Pa.

a. CA - 8-6 b. MA - 6-7

3. Family:

	Birthplace	Occupation	Education	Marital Status	No. Children
Father	New Castle	Laborer	Not HS Grad	Married and Living Together	3
Mother	New Castle	Housewife	Not HS Grad		

4. Test Results:

Intelligence Tests

	Name of Test	Date	Grade	Result
1.	Lee Clark Reading Readiness	5/64	Preschool	48 High
2.	Pintner-Cunningham	9/64	1st	79 IQ
3.	Binet	11/65	2nd	93 IQ

5. Reading Levels:

Year	Grade	Treatment	RE	RL*	Degree of Retardation
1964-65	1st	i/t/a-Merr	1.8	1.0	.8
1965-66	2nd	i/t/a-Merr	2.6	1.1	1.5
1966-67	2nd (R)	RC	3.4	1.7	1.7

* Stanford Achievement Test - Paragraph Meaning Subtest - administered at end of respective school years.

APPENDIX F (CONTINUED)

6. Diagnostic Considerations:

   a. Background Information:

      1964-65: Larry's mother has taken the attitude that her son is "stupid and can't learn." He has received special help from the teacher.

      1965-66: Larry is a discipline problem. He is immature and has poor social and emotional development.

      1966-67: Larry is receiving professional guidance for his emotional problem. He is often very hostile toward his peers. The guidance counselor has reported him to be in need of attention and affection.

   b. Intellectual:

      1964-65: Larry scored a high average on his reading readiness test which corresponds to a .7 grade level equivalent. His Pintner-Cunningham showed a 79 IQ which is classified as borderline defective.

      1965-66: On a Binet, Larry scored a 93 IQ which is classified as normal or average.

   c. Educational:

      1964-65: Larry's SAT paragraph meaning score showed him to be a nonreader.

      1965-66: Larry's SAT score still showed him to be a near nonreader.

      1966-67: Larry gained six months in his reading level and remained one year, seven months behind his reading expectancy score. This was a relative improvement for him.

APPENDIX F (CONTINUED)

Scott, Foresman-A

CASE STUDY OUTLINE

1. Name: Morris

2. Birthday: February 3, 1958
New Castle, Pa.

a. CA - 9-4 b. MA - 6-7

3. Family:

	Birthplace	Occupation	Education	Marital Status	No. of Children
Father	New Castle	Railroad	Not HS Grad	Separated	
Mother	New Castle	Housewife	Not HS Grad	Separated	2

4. Test Results:

Intelligence Tests

	Name of Test	Date	Grade	Result
1.	Lee Clark Reading Readiness	5/64	Preschool	31 Low
2.	Pintner-Cunningham	9/64	1st	72 IQ
3.	Pintner-Cunningham	11/65	1st (R)	89 IQ

5. Reading Levels:

Year	Grade	Treatment	RE	RL*	Degree of Retardation
1964-65	1st	SF	1.7	1.2	.5
1965-66	1st (R)	RC	2.4	1.5	.9
1966-67	2nd	RC	3.2	1.9	1.3

* Stanford Achievement Test - Paragraph Meaning Subtest - administered at end of respective school years.

6. Diagnostic Considerations:

   a. Background Information:

      1964-65: Morris was quite disturbed by his parents' separation. He is a careless worker and a constant talker.

   b. Intellectual:

      1964-65: Morris scored a low average on his preschool reading readiness test which corresponded with a .2 grade level equivalent. His Pintner-Cunningham showed a 72 IQ which is classified as borderline defective.

      1965-66: On another form of the Pintner-Cunningham, Morris scored an 89 IQ which is low average.

   c. Educational:

      1964-65: Morris' reading level was close to that of a nonreader and was five months behind his reading expectancy.

      1965-66: Morris gained three months in his reading level and remained nine months behind his reading expectancy.

      1966-67: Morris gained four months in his reading level and remained one year, three months behind his reading expectancy.

APPENDIX F (CONTINUED)

Phonics and Word Power-A

CASE STUDY OUTLINE

1. Name: Thomas

2. Birthday: December 20, 1958
Ohio

a. CA - 8-6   b. MA - 6-6

3. Family:

	Birthplace	Occupation	Education	Marital Status	No. of Children
Father	Ohio	Laborer	Not HS Grad	Married and Living Together	5
Mother	Pennsylvania	Housewife	Not HS Grad		

4. Test Results:

Intelligence Tests

	Name of Test	Date	Grade	Result
1.	Lee Clark Reading Readiness	5/64	Preschool	13 Low
2.	Pintner-Cunningham	9/64	1st	78 IQ
3.	Stanford Binet	10/64	1st	74 IQ
4.	Pintner-Cunningham	10/65	1st (R)	100 IQ

5. Reading Levels:

Year	Grade	Treatment	RE	RL*	Degree of Retardation
1964-65	1st	PWP	1.8	1.2	.6
1965-66	1st (R)	PWP	2.6	1.4	1.2
1966-67	2nd	PWP	3.3	2.0	1.3

* Stanford Achievement Test - Paragraph Meaning Subtest - administered at end of respective school years.

6. Diagnostic Considerations:

   a. Background Information:

      1964-65: It had been recommended that Thomas wait another year as he was just not ready to profit from first-grade work. He has a slight hearing problem and his vision has been corrected. His attendance is not very good.

      1965-66: Thomas would be able to do better if he would pay attention and be more careful in his work habits.

   b. Intellectual:

      1964-65: Thomas' low reading readiness score recommended that a delayed entry would be advisable. His Pintner-Cunningham showed a 78 IQ which is termed borderline defective. On a Binet, Thomas scored a 74 IQ which is also classified as borderline defective.

      1965-66: On another form of the Pintner-Cunningham, Thomas scored a 100 IQ which is classified as normal or average.

   c. Educational:

      1964-65: Thomas' reading level showed that very little reading ability had been achieved by him this year in first grade. His reading level was six months behind his reading expectancy level.

      1965-66: Thomas gained only two months reading level this year of his retention. There remained a one-year-two-month difference between his reading level and reading expectancy scores.

      1966-67: Thomas' reading level showed a six-month gain this year. There remained a one-year-three-month difference between his reading level and reading expectancy scores.

APPENDIX F (CONTINUED)

Phonics and Word Power-B

CASE STUDY OUTLINE

1. Name: Gregory

2. Birthday: August 30, 1958
   New Castle, Pa.

   a. CA - 8-9 b. MA - 6-4

3. Family:

	Birthplace	Occupation	Education	Marital Status	No. of Children
Father	New Castle	Laborer	Not HS Grad	Married and Living Together	2
Mother	New Castle	Housewife	Not HS Grad		

4. Test Results:

Intelligence Tests

Name of Test	Date	Grade	Result
1. Lee Clark Reading Readiness	5/64	Preschool	39 Low
2. Pintner-Cunningham	9/64	1st	73 IQ
3. Binet	9/65	2nd	109 IQ

5. Reading Levels:

Year	Grade	Treatment	RE	RL*	Degree of Retardation
1964-65	1st	PWP	1.7	1.1	.6
1965-66	2nd	PWP	2.5	1.7	.8
1966-67	2nd (R)	PWP	3.2	1.7	1.5

* Stanford Achievement Test - Paragraph Meaning Subtest - administered at end of respective school years.

APPENDIX F (CONTINUED)

6. Diagnostic Considerations:

   a. Background Information:

      1964-65: Gregory was absent twenty-five days this year. He is very immature. He was "placed" into second grade because of the reading study.

      1965-66: Gregory will not be passed this year.

   b. Intellectual:

      1964-65: Gregory scored a low average on his reading readiness test which corresponded to a .4 grade level equivalent. His Pintner-Cunningham showed a 73 IQ which is classified as borderline defective.

      1965-66: On a Binet, Gregory scored a 109 IQ which is classified as normal or average.

   c. Educational:

      1964-65: Gregory learned to read very little this year as his SAT score showed him to be a near nonreader. The difference between his reading level and reading expectancy was six months.

      1965-66: Gregory showed a six-month gain on his reading level, and remained eight months behind his reading expectancy.

      1966-67: Gregory's reading level remained the same this year and he remained one year, five months behind his reading expectancy.

Probably the most common element running through the teachers' comments regarding retention was the idea that failure is an individual problem brought on by many factors - both personal and academic. Therefore, no single educational program is likely to be appropriate for all pupils. That the Lippincott program did account for forty percent of all retainees suggests that for the potential retainee this was not the most beneficial approach. This conclusion is similar to the suggestion, included in the discussion section of the December 1966 report of the second year of the New Castle Reading Study (22), that the Lippincott teachers primarily used a whole-class approach. It was further suggested that perhaps through ability grouping and other methods of meeting individual differences, the large retention figure could have been reduced.

As shown by the average reading level gains, those pupils retained in second grade showed more relative improvement than did those retained in first grade (Tables 2, 3, and 4).

TABLE 2
TREATMENT GROUP COMPARISONS OF READING EXPECTANCY AND READING LEVEL SCORES
GROUP A*

		1964-65			1965-66			1966-67		
	%**	RE	RL	DR	RE	RL	DR	RE	RL	DR
Lipp	24%	1.8	1.2	.6	2.6	1.6	1.0	3.5	2.2	1.3
PWP	18%	1.8	1.4	.4	2.5	1.5	1.0	3.2	2.2	1.0
SF	7%	1.7	1.3	.4	2.5	1.8	.7	3.3	2.3	1.0
i/t/a-Merr	5%	1.8	1.4	.4	2.6	1.4	1.2	3.4	1.8	1.6

* Entered Grade 1 in 1964-65 and were retained in Grade 1. (See pages F-5 and F-6 of this Appendix for explanations of abbreviations.)

** Percentage of total number of retainees.

APPENDIX F (CONTINUED)

TABLE 3
TREATMENT GROUP COMPARISONS OF READING EXPECTANCY AND READING LEVEL SCORES
GROUP B*

	1964-65			1965-66				1966-67		
	RE	RL	DR	%**	RE	RL	DR	RE	RL	DR
Lipp	1.8	1.4	.4	9%	2.6	1.8	.8	3.4	2.7	.7
PWP	1.8	1.4	.4	5%	2.6	1.9	.7	3.5	2.8	.7
i/t/a-Merr	1.9	1.4	.5	7%	2.7	2.0	.7	3.6	2.4	1.2
SF	(There were no Scott, Foresman pupils retained in this sample group studied)									

* Entered Grade 1 in 1964-65 and were retained in Grade 2. (See pages F-5 and F-6 for explanation of abbreviations.)

** Percentage of total number of retainees.

TABLE 4
TREATMENT GROUP COMPARISONS OF READING EXPECTANCY AND READING LEVEL SCORES
GROUP C*

	1965-66				1966-67		
	%	RE	RL	DR	RE	RL	DR
Lipp	7%**	1.9	1.3	.6	2.8***		
PWP	5%	2.1	1.4	.7	3.1		
SF	2%	1.9	1.5	.4	2.9		
i/t/a-Merr	11%	1.8	1.3	.5	2.4		

* Entered Grade 1 in 1965-66 and were retained in Grade 1. (See pages F-5 and F-6 for explanation of abbreviations.)

** Percentage of total number of retainees.

*** First graders were not administered the SAT this year.

APPENDIX F (CONTINUED)

Although this cannot be attributed definitely to any specific reason, it may be that those retained in second grade had more of a foundation established upon which to build. Also, it may be true that retained second graders were more mature and thus better qualified to handle academic demands. Saying it another way, perhaps those retained in first grade spent much of the second year in grade one gaining social maturity rather than improving in academic achievement.

As can be seen by the case study examples, the retainees are those children whose individual needs were not met. They are representative of the minority who require the individual diagnosis, attention and personalized instruction not provided for in the classroom. The average IQ reading expectancy and reading level figures of the retainees do not show any great differences among or between the treatment groups. The subjects are the same; the treatments are different. However, whether it can be said that the Lippincott group had forty percent of the total retainees because of its whole-classroom approach cannot be known by this case study review. There are other outside factors to consider here. The school with the greatest number of Lippincott retainees is located in a very low socio-economic area. The Lippincott teacher with the most retainees was one who set certain standards which had to be attained before promotion could occur. The involved school also was the least yielding to any outside advice concerning retentions.

Exactly how much influence each factor had cannot be ascertained precisely. All the factors, however, add up to a situation in which many have failed.